Role of Hel308 helicase in origin-independent DNA replication in two species of Haloferax

Catherine Imogen Harrison, BSc

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

March 2024

Abstract

DNA replication is a process which is essential to all forms of life. This process is initiated at defined points, known as origins of replication. Origins are universal, and were previously believed to be essential. However, the halophilic euryarchaeon *Haloferax volcanii* is able to survive in the absence of origins, and in fact, origin-deleted mutants grow faster than the wild-type. Growth in these circumstances is thought to be dependent on recombination-dependent replication, since the recombinase protein RadA is essential in origin-deleted cells.

The Superfamily 2 DNA helicase Hel308 plays a similar role to RecQ family helicases, and has been implicated in genome stability and DNA repair. Like RecQ, Hel308 has a preference for ssDNA and forked structures *in vitro*. The *hel308* gene is inessential in *H. volcanii*, and its deletion is associated with growth defects and increased susceptibility to DNA damage. ATPase-null point mutants of this protein have been explored previously, with one (K53A) reported to be lethal in both *H. volcanii* and *E coli*.

In the closely related species *Haloferax mediterranei*, origins have been shown to be essential, in contrast to *H. volcanii*. It was also previously reported that the *hel308* gene from this species was essential. A genetic screen previously identified this gene as a potential inhibitor of origin-independent replication in *H. volcanii.*

In the work presented here, *hel308* and its relevance to origin-independent replication were explored. It was found that *H. mediterranei hel308* does not prevent origin-independent replication in *H. volcanii*, either through a system with an inducible origin, or through inducible expression of *hel308* in a strain lacking origins.

The ATPase-null point mutant was not found to be lethal in *H. volcanii*, and could be engineered onto the chromosome. Brief exploration of this strain revealed a phenotype similar to both *Δhel308* and previous ATPase-null point mutants, but not exactly equivalent.

Attempts were made to identify the factors contributing to the essentiality of *hel308* in *H. mediterranei*, which questioned the original report of essentiality of this gene. Despite further investigation, essentially of *hel308* in *H. mediterranei* has not been conclusively proven or disproven.

Origin-deleted cells also show increased resistance to the family B DNA polymerase inhibitor aphidicolin, suggesting a differential usage of DNA polymerases in these cells. Meanwhile, *hel308*-deleted cells show increased susceptibility to aphidicolin. Responses to aphidicolin were explored in *H. volcanii* strains with and without origins, and with various *hel308* alleles present or absent. The susceptibility to aphidicolin conferred by *hel308* deletion appears dominant to the aphidicolin resistance seen in originless strains. However, very low doses of aphidicolin induced a small increase in growth rate in strains lacking origins and with functional Hel308 present. It is suggested that these effects may be due to differential DNA polymerase usage and the importance of Hel308 in processing stalled forks following aphidicolin treatment. Preliminary results concerning the response of *H. mediterranei* to aphidicolin are intriguing, but do not shed conclusive light on its workings.

The interplay between origins, recombination and repair are complex, and warrant further work. Differential usage of DNA polymerases in the presence or absence of origins could provide important insight into the mechanisms of origin-initiated and recombination-dependent replication in archaea.

Acknowledgements

I acknowledge nothing and I never will.

That said, there are many people who need to be thanked for their support, advice and patience during the last 4.5 years.

Firstly Thorsten, for his inexhaustible advice and unfailing optimism in the face of inconvenient results. His insights and anecdotes proved invaluable in my battle with *H med*.

Laura, for her patience and indulgence in my ongoing quest to locate anything in the lab.

Ambika, for passing me the (cursed) *H med* baton, and for helping me settle in.

Fellow lab-lings Ryan, Kate, Andy and Anna for their camaraderie and commiserations when the data were misbehaving.

And of course, Josh, for his ongoing support and affection when the salty boys were getting me down. His excellence is most thoroughly acknowledged.

Abbreviations

5-FOA	5-fluoroorotic acid
AAA+	ATPase associated with diverse cellular activities
ADP	Adenosine 5’-diphosphate
Amp	Ampicillin
ATP	Adenosine 5’triphosphate
BER	Base excision repair
BIR	Break induced replication
bp	Base pair(s)
Cas	Casamino acid-based media
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
D-loop	Displacement loop
dHJ	Double Holliday junction
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic acid
DNAP	DNA polymerase
dNTP	Deoxynucleotide
DPBB	Two-double-psi-ß-barrel
DSB	Double-strand (DNA) break
DUE	Duplex unwinding element
EDTA	Ethylenediaminetetraacetic acid
GFP	Green fluorescent protein
HJ	Holliday junction
His	Histidine
HR	Homologous recombination
HTH	Helix-turn-helix
IDR	Intrinsically disordered region
kb	Kilobase(s)
LB	Lysogeny broth
Leu	Leucine
Mb	Megabase(s)
MMC	Mitomycin C
MMEJ	Microhomology-mediated end joining
MMR	Mismatch repair
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
OD	Optical density
OIR	Origin-independent replication
ORB	Origin recognition box
ORC	Origin recognition complex
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIP	PCNA-interaction motif
R-loop	RNA-based displacement loop
RDR	Recombination-dependent replication
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SAMP	Small archaeal modifier protein
SDS	Sodium dodecyl sulfate
ssDNA	Single-stranded DNA
SSPE	Saline sodium phosphate EDTA

SW	Salt water
TAE	Tris acetic acid EDTA
TBE	Tris boric acid EDTA
Thy	Thymidine
TLS	Translesion synthesis
Trp	Tryptophan
US	Upstream
Ura	Uracil
UV	Ultraviolet light
v/v	Volume per volume
w/v	Weight by volume
WH	Winged helix
WT	Wild type
X-gal	5-bromo-4-chloro-3-indolyl-ß-D-galactopyranoside
YPC	Yeast extract, peptone, casamino acid-based media

Contents

1. Introduction

1.1 Archaea

1.1.1 Discovery of archaea

Archaea are a domain of single-celled organisms distinct from both bacteria and eukaryotes. Prior to 1977 what are now known as archaea were not distinguished from other bacteria in the prokaryotic domain. It was the work of Woese and Fox that suggested the reclassification of these organisms into a new, separate 'urkingdom', termed 'archaebacteria' (Woese & Fox, 1977). Woese and Fox's analysis was based on comparative sequence analysis of the 16S rRNA common to all life forms, and resulted in a reorganisation of the tree of life, firstly into three domains; bacteria, archaea and eukarya. Further analysis has revealed that archaea are much more similar to eukarya than they are bacteria, despite their prokaryotic morphology (Woese et al., 1990).

Like bacteria, archaea are single-celled organisms that lack a nucleus to contain their circular chromosomes, where their relatively small genomes are frequently organised into operons (Santangelo et al., 2008). However, their internal workings are much more complex than their bacterial equivalents, appearing more like simplified versions of the eukaryotic systems in many cases. For example, archaea have been found that contain histone-equivalent proteins (Brunk & Martin, 2019), a ubiquitin-like system (SAMP) (Darwin & Hofmann, 2010), actin-like cytoskeletons (Albers et al., 2022) and ESCRT-like membrane remodelling machinery (Hatano et al., 2022). A summary of the prevalence of these features in different families of the archaea are shown in Figure 1, below.

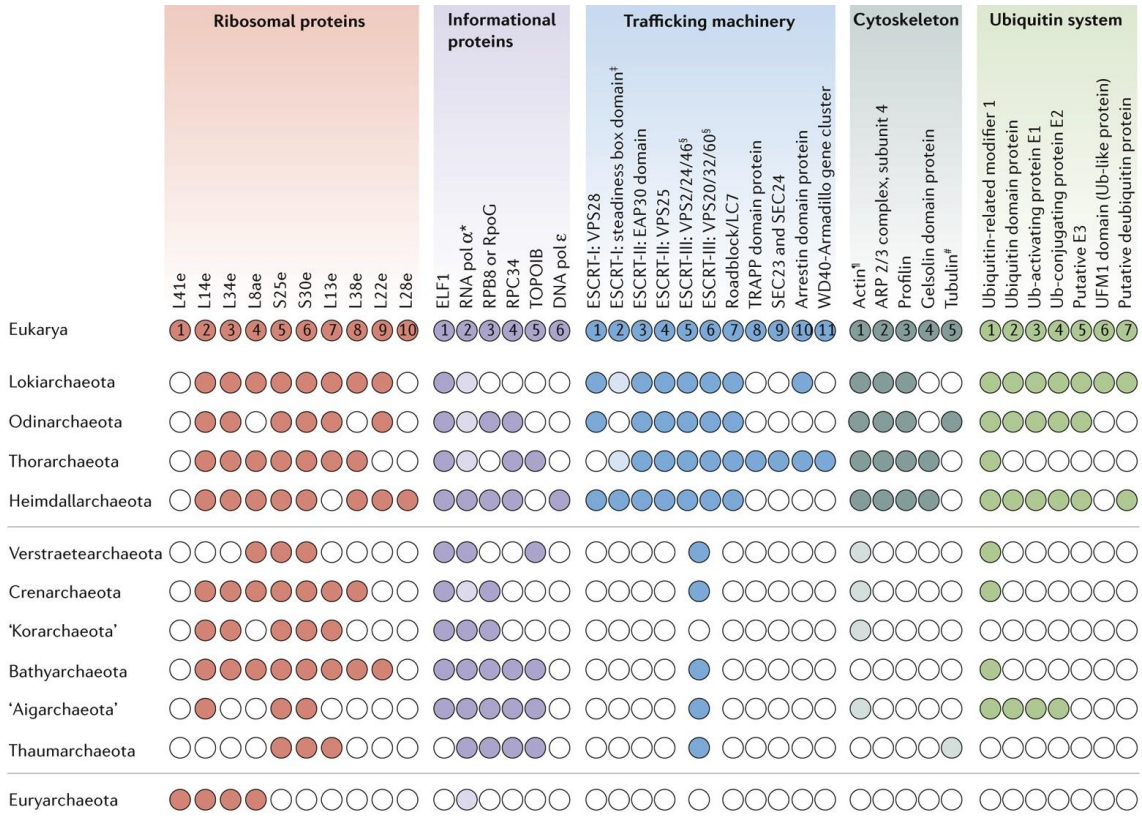


Figure 1 Prevalence of eukaryotic signature proteins in different archaea lineages. From (Eme et al., 2017)

It has since been proposed that the eukarya and archaea are either sister groups sharing a common ancestor, or that ancestors of the eukarya have arisen directly from the archaea. These will be discussed below.

Archaea are often considered extremophiles, having been found all over the world in a variety of seemingly inhospitable habitats, including extremes of cold, heat, acidity, alkalinity, salinity, and pressure (Merino et al., 2019). However, improvements in sampling techniques have revealed archaea in more mundane, mesophilic habitats, such as soil and seawater; some have even been found in the human body (Nguyen-Hieu et al., 2012; Thomas et al., 2022). It seems that, while many extremophiles are archaea, far from all archaea are extremophiles.

In addition to their relevance to exploring the evolution of eukaryotes, archaea are often studied for practical applications. For example, DNA polymerase enzymes isolated from hyperthermophiles are now in use in PCR applications (Ishino & Ishino, 2014). Archaea may also represent an as-yet mostly untapped source of novel antimicrobial agents. Several peptides with antimicrobial properties have been isolated from members of the Euryarchaea and Crenarchaea, collectively termed archaeocins (Besse et al., 2015). These are divided into two classes; halocins (from the *Halobacteriales*) and sulfobiocins (from the *Sulfolobales*). While most of those known seem to target related species of archaea, some have been demonstrated to inhibit a variety of bacteria, including some pathogenic species (Kavitha et al., 2011).

There is also interest in usage of archaea in industrial applications. Archaea are a metabolically diverse group, and are able to produce a range of potentially useful metabolites and chemicals. For example, all methanogens are archaea; therefore all methane produced for biofuel can only be sourced by archaea (Enzmann et al., 2018). Halophiles also produce polyhydroxyalkanoates as a storage polymer, which can be harvested to produce bioplastics as an alternative to fossil carbon sources (Mitra et al., 2020).

1.1.2 Taxonomy/phylogeny of archaea

Further rRNA analysis first divided the archaea into two main phyla; the Crenarchaeota and the Euryarchaeota (Winker & Woese, 1991). Since this time, new species have been identified frequently, and sampling techniques have advanced considerably, allowing identification of novel archaea species from traces of their genomic DNA retrieved from their various habitats (Spang et al., 2015). This has greatly increased the number of available genomes for phylogenetic analyses, and has necessitated several re-organisations in the tree of life.

One of these reorganisations is the recognition of the TACK superphylum, composed of the Thaumarchaeota, Aigarchaeota, Crenarchaeota and Korarchaeota (Guy & Ettema, 2011). TACK is now considered something of a sister group to the euryarchaeota. Some other rearrangements have been the sprouting of new branches of the tree of life, such as the discovery of the Nanoarchaeota, studied for their highly concise genome and small cell size (Wurch et al., 2016). These have since been placed in the DPANN superphylum, made up of the Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota and Nanohaloarchaeota (reviewed in Dombrowski et al. (2019)). Since its proposal, the DPANN superphylum has been bolstered by the addition of other phyla, including the Woesearchaeota, Pacearchaeota and Altiarchaeota.

DNA samples taken from marine sediments near deep-sea hydrothermal vents provided another branch to the tree in the form of the Lokiarchaeota (Spang et al., 2015). Further samples have since identified a range of other clades, including the Odinarchaeota, Thorarchaeota and others, which form the varied Asgard superphylum. Many analyses have placed Asgard archaea as the closest

ancestor to the eukaryotes; specifically as a sister group to the Hodarchaea, within the group Heimdallarchaeia (Eme et al., 2017). Recent culturing of an organism from the Lokiarchaeota for the first time revealed that it contains hundreds of genes that are considered eukaryotic signatures – including four proteins homologous to eukaryotic actin (Rodrigues-Oliveira et al., 2023).

One barrier to further examination of the Asgard superphylum is that many of its members are only attested through trace environmental DNA, and the host organism has never been seen nor cultured. However, it has been shown that some of these species can be cultured, but only in co-culture with other organisms, perhaps suggesting an obligate symbiosis (Rafiq et al., 2023). Others are extremely slow growing – some requiring culturing for over 2000 days for detectable growth (Aoki et al., 2014). Recent cultivation of the Asgard organism *Prometheoarcheum syntrophicum* further supported the close evolutionary relationship between this superphylum and the eukarya. This species bears a high number of genes which demonstrate similarity to eukaryotic mechanisms (Imachi et al., 2020). Furthermore, the complex stalks and branching structures of their cells are suggestive of a potential pathway towards endosymbiosis, where the branches might capture bacterial cells and slowly engulf them. It is proposed that this process allowed the eukaryotic ancestor to capture the ancestors of the mitochondrion; an important step in eukaryogenesis (Baum & Baum, 2014).

Archaea are therefore of considerable utility in exploring the ancestry and evolution of eukaryotic cellular mechanisms. Their internal workings are often revealed to be simplified versions of their eukaryotic equivalents, and their single-celled nature makes them amenable to a range of adapted techniques well-established in bacteria. Several species of archaea are well-studied model organisms with well-developed toolkits available for their manipulation, reviewed briefly in Harrison & Allers (2022).

1.2 Haloferax volcanii

1.2.1 Overview

Haloferax volcanii is a halophilic euryarchaeon originally isolated from the Dead Sea in 1975 (Mullakhanbhai & Larsen, 1975). It is an extremophile, growing optimally in 1.7-2.5M NaCl, although tolerating up to 3.5M (Bidle et al., 2008). It is amenable to culturing in the lab, being a chemoorganotroph, capable of utilising a range of sugars, polysaccharides and amino acids as a carbon source (Pohlschroder & Schulze, 2019). It grows optimally at 45°C, with a typical doubling time of 2-3 hours. Cells are usually coccoid, but can vary in morphology greatly depending on conditions (de Silva et al., 2021). Carotenoids and bacteriorubin present in the cell lend a distinctive red colour, both in broth and in colonies; these chemicals are thought to help protect the cell from oxidative damage (Giani & Martínez-Espinosa, 2020). Like many archaea, *H. volcanii* is surrounded by a glycoprotein S-layer, as opposed to the peptidoglycan cell walls of bacteria (von Kügelgen et al., 2021).

Many salt-tolerant bacteria have adapted to their high-salt environments by adopting a “salt-out” mechanism. These organisms maintain osmotic homeostasis by pumping ions out of the cell, and compensating for the resulting osmotic gradient by packing organic solutes into the cell, such as sugars, amino acids and modified derivatives of these (Shivanand & Mugeraya, 2011). In contrast, *H. volcanii* adopts a “salt-in” mechanism, wherein the intracellular environment is maintained at a high ionic concentration roughly equivalent to that of the environment. This is usually achieved through accumulation of potassium cations. This necessitates adaptation of the cellular machinery to the

high-salt internal environment. For this reason, the genomes of halophilic organisms are often higher in GC content, conferring additional stability to the duplex due to the stronger interaction between these bases (Paul et al., 2008). Halophilic proteins often bear a high proportion of acidic amino acids on their external surfaces, as interactions between the negatively charged residues and the ions in solution encourage solubility (Allers 2010). As a result, halophilic proteins have a tendency to misfold and precipitate from solution in low ionic strength environments, making heterologous expression difficult.

1.2.2 Genome structure

Archaea contain circular replicons, much like bacteria. Depending on their qualities, these can be categorised as chromosomes, plasmids, or mini-chromosomes/ megaplasmids. Plasmids are small and dispensable, bearing no essential genes. They are also not necessarily native to the host, being horizontally transferred between species with relative frequency in their natural habitat (Chimileski et al., 2014). Mini-chromosomes/ megaplasmids are larger and may bear essential genes, making them indispensable to the host. Wild isolates of *Haloferax volcanii* contain one large circular chromosome (2.8 Mb), one plasmid; pHV2 (6kb), and two megaplasmids; pHV3 (438 kb), and pHV4 (636 kb). The remaining replicon, pHV1 (85kb), has been referred to as either a plasmid or a megaplasmid, based on its middling size, non-native origin and inessentiality.

In the lab strain and its descendants, the small pHV2 plasmid has been intentionally cured so that plasmids may be constructed using its origin (Wendoloski et al., 2001). The mini-chromosome pHV4 has also stably incorporated into the main chromosome, resulting in a 3.5Mb main chromosome (Hawkins et al., 2013); see Figure 2 (page 9).

Haloferax is highly polyploid, with *H. volcanii* maintaining around 20 copies of the main chromosome per cell, and *H. mediterranei* higher still, although copy number in both species is influenced by growth conditions and nutrient availability (Bruert et al., 2006). In particular, increased phosphate levels result in increased copy number, suggesting that additional genome copies may function as a phosphate storage mechanism (Zerulla et al., 2014).

1.2.3 Genetic toolkit

Despite its extremophilic lifestyle, *H. volcanii* is surprisingly amenable to culturing under laboratory conditions. Decades of work with this organism have resulted in a wide range of tools for its genetic manipulation.

Culturing of *H. volcanii* is routinely carried out in rich medium containing yeast extract, peptone and casamino acids (YPC) (Allers et al., 2010). However, selective media based on casamino acids (Cas) is also viable (de Silva et al., 2021). The benefit of this media is that it contains very low levels of nucleotides, tryptophan and leucine; all of which can be exploited for selection of different commonly used marker genes. Minimal media is also available, to allow selection of more unusual genetic markers (described in Dattani, Harrison, et al., 2022).

1.2.4 Transformation and plasmids

While some archaea, such as *Thermococcus,* will take up free DNA from the environment (Matsumi et al., 2007), in *H. volcanii* the glycoprotein S-layer acts as an effective barrier to this natural competence. Transformation protocols for *H. volcanii* address this by first stripping the S-layer using EDTA, leaving mechanically fragile spheroplasts behind. Administering polyethylene glycol (PEG-600) then aids uptake of DNA through the exposed cell membrane (Cline et al., 1989).

H. volcanii wild isolates include several autonomously replicating sequences, including the small pHV2. This naturally occurring plasmid has been intentionally cured from the lab strain, allowing its origin to be repurposed for use in plasmids in this background (Wendoloski et al., 2001). Several other origins have also been identified for use in plasmids, with varying qualities in terms of their copy number and persistence in the cell (Norais et al., 2007).

Haloferax species bear a restriction endonuclease (Mrr) that degrades methylated DNA, as a defence mechanism against foreign genetic material and viruses (Allers et al., 2010). This greatly reduces the efficiency of transformation with methylated DNA; for example, plasmids that have been prepared in many standard lab strains of *E. coli* bear *dam* methylation patterns (Marinus, 2000). This can be avoided by passaging plasmids in *dam- E. coli* strains, but these strains are subject to a higher level of mutation as a result of interactions with their DNA repair machinery. Both problems can however be avoided by deletion of the *mrr* gene (encoding the Mrr restriction enzyme), allowing good transformation efficiency with methylated DNA (Allers et al., 2010).

In addition to transformation with prepared DNA, *Haloferax* is capable of genetic exchange between members of its population through cytoplasmic bridges (Rosenshine et al., 1989). This is often referred to as mating, and can be a useful tool in the lab to combine genotypes and qualities of different strains. In addition to horizontal gene transfer within the species, *Haloferax volcanii* can also undergo cell fusion with the related species, *H. mediterranei*, to produce hybrid cells (Naor, et al., 2012).

1.2.4.1 Selectable markers

While many bacterial model systems make use of antibiotic resistance genes, for a long time these were problematic in *H. volcanii* due to issues with recombination between the plasmid-borne resistance markers and the wild-type allele on the host chromosome. The resistance genes for mevinolin (*hmgA*) and novobiocin (*gyrB*) were initially identified through *H. volcanii* mutation screens (Holmes & Dyall-Smith, 1991; Lam & Doolittle, 1989). The mutant alleles were thus highly similar to the chromosomal wild-type allele, which could not be deleted due to its essentiality. Plasmids bearing these marker genes would frequently recombine onto the main chromosome at the site of the wild-type allele. These issues were eventually circumvented though the use of homologues from closely related species, to reduce allele similarity and prevent these recombination events (Wendoloski et al., 2001).

Many of the more commonly used selectable markers in *H. volcanii* are genes involved in essential amino acid or nucleotide biosynthesis pathways. With the corresponding gene deleted from the genome, presence of a plasmid bearing the complementing gene can be detected by restored autotrophy for the specific component. These markers are typically paired with the promoter from *Halobacterium salinarum*'s ferrodoxin gene to ensure strong expression (Gregor & Pfeifer, 2005). The most common biosynthesis marker genes are summarised in the table below. Of these, the most frequently used in the lab are *pyrE2, trpA,* and *hdrB.*

Table 1-1 Common selectable biosynthesis markers used in the study of Haloferax volcanii*.*

Gene	Enzyme name	Pathway	Reference
pyrE2	Oronate phosphoribosyl transferase	Uracil biosynthesis	(Bitan-Banin et al., 2003)
trpA	Tryptophan synthase	Tryptophan biosynthesis	(Allers et al., 2004)
leuB	3-isopropylmalate dehydrogenase	Leucine biosynthesis	(Allers et al., 2004)

hdrB	Dihydrofolate reductase	Thymidine biosynthesis	(Ortenberg et al., 2000)
hisC	Histidinol-phosphate aminotransferase	Histidine biosynthesis	(Leigh et al., 2011)
metX	Homoserine O-acetyltransferase	Methionine biosynthesis	(Leigh et al., 2011)
argH	Argininosuccinate lyase	Arginine biosynthesis	(McMillan et al., 2018)
lysA	Diaminopimelate decarboxylase	Lysine biosynthesis	(McMillan et al., 2018)

The *pyrE2* marker gene is of particular utility in genetic editing applications, as there are protocols for both selection and counter-selection using this marker (Dattani, Harrison, et al., 2022). *pyrE2* is necessary for the uracil biosynthesis pathway; cells bearing functional copies of this gene are uracil autotrophs, and can be selected for through their ability to grow on media lacking this component. However, the enzyme encoded by this gene will also break down 5-fluoroorotic acid (5-FOA) to 5-fluorouracil, which is toxic to the cell. Strains in which this gene has been deleted are therefore resistant to 5-FOA, while cells carrying this gene will rapidly accumulate toxin and die; thereby allowing efficient counter-selection against *pyrE2* (Allers et al., 2004).

1.2.4.2 Reporter genes

While not directly selectable, reporter genes can provide valuable information as to the genetic composition of a strain. Some well-known marker genes have been adapted from other systems to work in the high-salt environment of *H. volcanii*'s cells.

H. volcanii naturally lacks β-galactosidase activity, and so cannot cleave X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) to result in production of the blue 5,5'-dibromo-4,4'-dichloro-indigo. The *E. coli* reporter gene *lacZ* cannot be used in *Haloferax* due to the high salt causing misfolding of the protein; however, an effective homologue has been isolated from *Haloferax alicantei*, known as *bgaH* (Holmes & Dyall-Smith, 2000). This gene is functional in *H. volcanii* and allows screening of colonies by spraying them with a solution of X-gal and allowing the blue colouring to develop. Given the natural colouring of *Haloferax*, this is referred to as blue-red screening.

Salt-tolerant GFPs have also been developed for use in *Haloferax* (Reuter & Maupin-Furlow, 2004). In addition to their usual suite of applications in cell biology, these have also been developed into an interaction assay, similar to yeast two-hybrid analyses. The GFP gene is split into two fragments; NGFP and CGFP. These fragments are then fused to other genes to produce proteins each bearing one half of the GFP protein. The two halves of the GFP do not fluoresce when produced *in trans*, but close interactions between the proteins they are fused to can bring the two halves together, restoring fluorescence and effectively reporting the interaction between the proteins carrying them (Winter et al., 2018). This is useful as the halophilic modifications of many *Haloferax* proteins preclude their investigation through yeast two-hybrid analyses, as they are liable to misfold and become inactive in the intracellular environment of *Saccharomyces*.

The biosynthesis pathway that produces carotenoid pigments in *Haloferax volcanii* can also be disrupted without causing a growth defect to the cells (Turkowyd et al., 2020). This is achieved through deletion of the *crt1* gene, which encodes a phytoene dehydrogenase necessary for lycopene synthesis, which is a precursor compound in the carotenoid synthesis pathway. Deletion of this gene results in *Haloferax* colonies that appear white rather than red. While not a marker gene *per se*, this

can be used to differentiate different cell populations or allele presence; for example, in Dattani, Sharon, et al. (2022). This modification also reduces autofluorescence in *Haloferax* cells, which is convenient for imaging techniques using other fluorescent markers (Turkowyd et al., 2020).

1.2.4.3 Expression control

Several tools are available for the control of gene expression in *H. volcanii* cells. The promoter region taken from *Halobacterium salinarum*'s ferrodoxin gene (*p.fdx*) offers strong promotion of its downstream genes (Gregor & Pfeifer, 2005). This is typically used to control expression of selectable marker genes, to ensure expression and selectivity.

The tryptophan-inducible promoter is well characterised and frequently used (Large et al., 2007). In the wild type, this promoter controls expression of a tryptophanase gene, which cleaves excess tryptophan into indole, pyruvate and ammonia (Pfeiffer & Dyall-Smith, 2021). This gene is not needed during times of tryptophan autotrophy, hence its tight repression in the absence of environmental tryptophan. In the lab, this promoter is often placed downstream of a terminator sequence (Shimmin & Dennis, 1996), which prevents through-transcription from other sites of transcription; ensuring that expression is solely dependent on tryptophan. In addition to the wild-type *p.tnaA* promoter, a mutant form of this promoter also allows inducible expression through tryptophan, but to a much lower maximum level of expression (Braun et al., 2019). Interestingly, the reduction in expression levels is caused by a single base difference in the promoter sequence.

More recently, a second inducible promoter has been identified; *p.xyl*, the xylose-inducible promoter (Rados et al., 2023). In combination with the tryptophan-inducible promoter, this allows independent control of expression of two different genes of interest within a single cell for the first time in this model organism.

Repression via CRISPR interference has also been demonstrated in *H. volcanii* (Stachler & Marchfelder, 2016). It has been demonstrated to reduce transcript levels and expression down to 8% (22% for an essential gene). The details of this system in *H. volcanii* are summarised in Maier et al. (2019).

1.2.4.4 Gene deletion/replacement

Genetic alterations of *H. volcanii* are typically carried out through the pop-in/pop-out system, which makes effective use of the selection and counterselection qualities of the *pyrE2* marker gene (Bitan-Banin et al., 2003). The details of this method are discussed in the Materials and Methods chapter, (Chapter 2).

While the theories underpinning the pop-in/pop-out method are relatively simple, the high ploidy of this genus provides a complicating factor, in that the intended modification may not be present on all copies of the chromosome (Dattani, Harrison, et al., 2022). These strains are referred to as "merodiploid" and are typically unstable. *H. volcanii* undergoes frequent recombination between the many copies of its chromosomes, meaning that, in the absence of selective pressure, either of the alleles present could rapidly reach fixation within the population (Dattani, Sharon, et al., 2022). Homoploidy should therefore be verified before the genomic editing is considered complete; this is typically performed through methods such as Southern blot, which are highly sensitive.

1.3 DNA replication

1.3.1 DNA replication in *Haloferax*

Haloferax species have not been shown to demonstrate a distinct S-phase of the cell cycle; origins have varying levels of activity, and, while subject to various regulatory mechanisms, do not appear to be synchronised across the many copies of the chromosome within the cell (Hawkins et al., 2013). Deep sequencing of *H. volcanii* has shown that chromosomal regions near the most active origins can be represented at ratios greater than 2:1 compared to other sections of the genome, suggesting concurrent rounds of replication from these origins (Hawkins et al., 2013). Indeed, DNA replication seems to be largely uncoupled from the cell cycle; prevention of cell division results in accumulation of genome copies, with a ploidy of over 2000 inferred in some cases (Liao et al., 2021).

Within the Archaea, the number of origins per chromosome does not appear to be correlated with genome size (Ausiannikava and Allers 2017). The common lab strain of *H. volcanii* bears four origins on its circular chromosome, although one of these belongs to the incorporated pHV4, which bears its own origin of replication (ori-pHV4). The replicons and their origins are shown in Figure 2, below. The other *H. volcanii* mini-chromosomes bear their own origins of replication, whose differing levels of activity influence the copy number of these chromosomes (Norais et al., 2007). Origin sequences in *H. volcanii* consist of an AT-rich stretch of around 100 nucleotides (the duplex unwinding element; DUE), flanked by inverted origin recognition box (ORB) sequences (Pérez-Arnaiz et al., 2020). Initiation of replication at the origin is carried out by cdc6/orc1 proteins, so called because they show similarity to both a subunit of the eukaryotic origin recognition complex (ORC), and the cdc6 eukaryotic replicative helicase loader, and may be an ancestor of both eukaryotic proteins. These genes and their proteins will be referred to as *orc*/Orc for convenience.

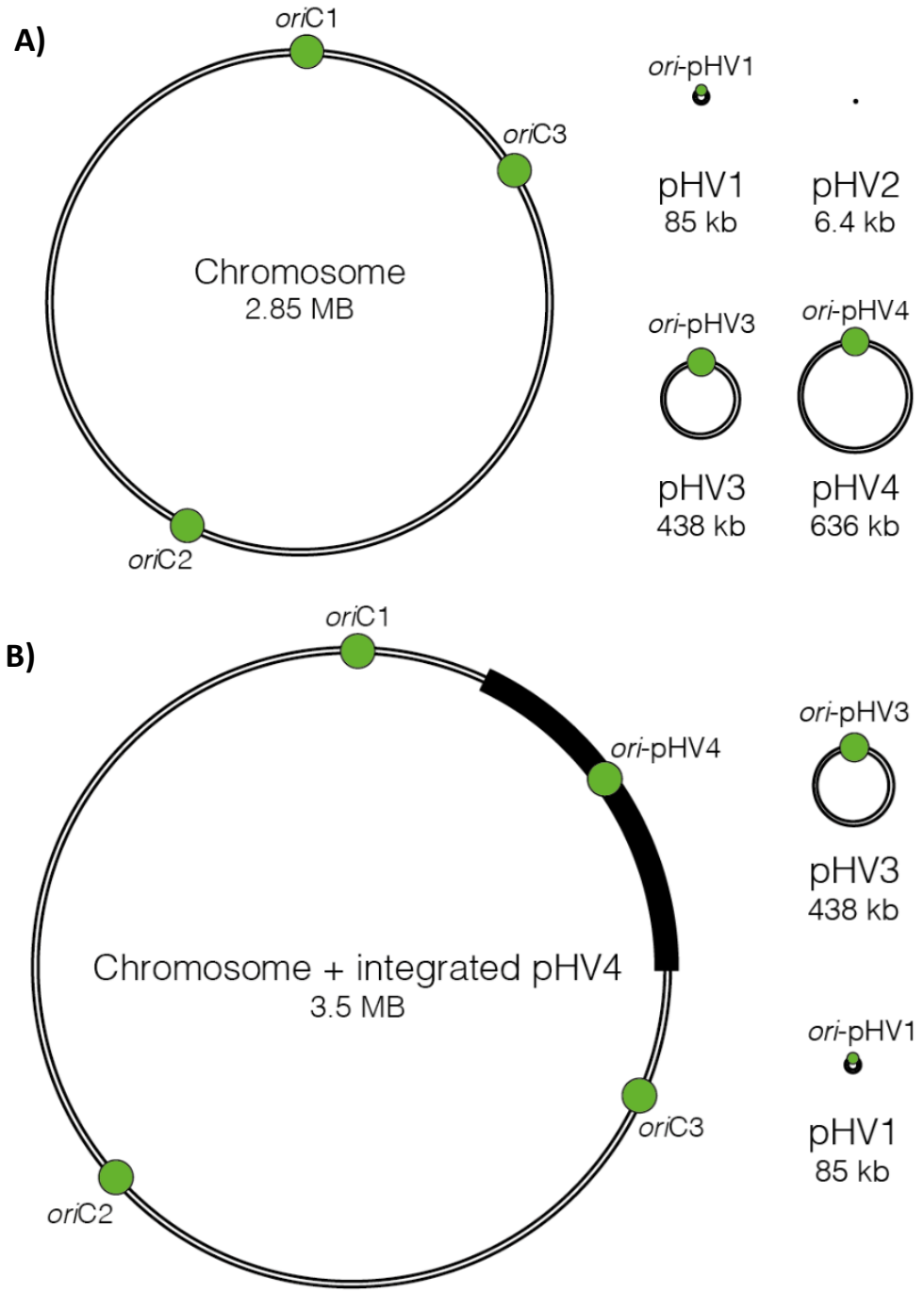


Figure 2. Genome structure of Haloferax volcanii*, showing the locations of DNA replication origins. (A) Genome structure of wild-type strain DS2, with a single circular chromosome, three mega-plasmids (pHV1, pHV3 and pHV4), and a small plasmid (pHV2). (B) Genome structure of laboratory strain H26 consisting of a single circular chromosome with the integrated pHV4 mega-plasmid (solid black line), and the mega-plasmids pHV1 and pHV3. From Marriott (2018).*

Haloferax origins each have an *orc* gene in close proximity. As archaea can perform synchronous transcription and translation, this may help ensure that the Orc protein will be produced close to its target origin (Kelman & Kelman, 2014). This could aid independent control of origins; however, chromosomal proximity of the gene is not necessary for origin activation (Norais et al., 2007). *H. volcanii* contains many more *orc* genes than origins, with sixteen *orc* genes identified in the genome, compared to six origins in the lab strain (Hartman et al., 2010).

These Orc proteins demonstrate varying activity, with some showing single-origin specificity, and some able to activate multiple origins (Ausiannikava, unpublished data). However, it is postulated that some of these Orc proteins are non-functional in the cell, possibly the result of orphan sequences whose origins are no longer present; these may have been acquired through duplication events or horizontal gene transfer (Pérez-Arnaiz et al., 2020). Differential translation and activity of these many Orc proteins plays a regulatory role in control of copy number of the cell's different replisomes (Babski et al., 2016). The one exception to this system is the pHV2 origin, which does not have an adjacent *orc* gene, and is not activated by an Orc protein. It is thought to replicate via a recombination-based method (Norais et al., 2007; Woods & Dyall-Smith, 1997). This origin maintains its replicon at ~6 copies per chromosome copy (Haque et al., 2020).

The evolutionary history of origins within the archaea seems to be a complicated story. Analysis of predicted origins and Orc proteins across a range of available archaeal genomes revealed that each haloachaeal species examined contained a copy of OriC1, which was likely inherited from the last archaeal common ancestor (LACA) (Wu et al., 2012). Other origins showed considerable diversity, and a pattern of heredity consistent with horizonal gene transfer, suggesting that they did not arise from duplication events. Even closely related species sometimes demonstrate markedly different genome structure and origin variety (Wu et al., 2012). This could explain why origins typically have their respective *orc* gene in close proximity; this ensures that they will be inherited together and remain functional in the event of transfer to another species. There is an argument that replicons within the same organism could be considered as evolving independently (Ausiannikava & Allers, 2017).

In addition to frequent horizontal transmission, *H. volcanii* shows a tendency towards genome rearrangements, and has been observed forming new, stable chromosomes and mini-chromosomes in the lab (Ausiannikava et al., 2018). It is likely that examination of wild populations may yield a range of minichromosomes, plasmids and other replicons not observed in the lab strain.

1.3.1.1 Initiation of replication

Initiation of replication begins with the binding of Orc proteins at the ORB sequences. The polarity of the ORB sequences determines the orientation of Orc protein binding; as ORBs often appear as inverted repeats around the DUE, they result in Orcs bound to the DNA with opposite orientation on either side of the DUE. In the haloarchaea, ORBs include an additional string of guanine nucleotides known as the G-string, which has been shown to enhance Orc binding in *Haloarcula hispanica* (Wu et al., 2014).

In *Haloferax*, the replicative helicase is MCM (minichromosome maintenance). This is an essential gene, and unlike its homologues in the eukarya, *H. volcanii* contains a single *MCM* gene, whose protein functions as a homohexamer. However, some archaeal species encode several MCM proteins, which may act as a heterohexamer, similar to its eukaryotic equivalent (Kristensen et al., 2014). MCM monomers are members of the AAA+ ATPase superfamily, possessing a noncatalytic N-terminal domain (containing a zinc-binding domain), a central AAA+ ATPase domain, and a winged helix-turn-helix (wHTH) domain at the end of the C terminal. The MCM hexamer possesses 3'-5' helicase activity once assembled, and double MCM hexamers have been shown to increase helicase activity (Chong et al., 2000).

1.3.1.2 CMG complex formation

ATP-bound Orc proteins can load a MCM hexamer onto the leading strand of the DNA, relative to their position (Samson et al., 2015). Studies in *Solfataricus islandicus* have shown that ATP-bound Orc readily recruits MCM helicase, but once the ATP bound to the Orc is hydrolysed, the ADP

remains securely bound to the protein, effectively inactivating it and preventing additional helicase recruitment (Samson et al., 2013). In Eukarya, the next step in replication initiation is the formation of the CMG complex (Cdc45, MCM, GINS). In the Eukarya, GINS is a heterotetrameric complex composed of four subunits: Sld5, Psf1, Psf2 and Psf3; GINS is an acronym derived from the Japanese names of the numbers included in these subunit names. In the archaea, however, the GINS homologue varies by species. In some archaea, the GINS complex is composed of two proteins (GINS51, roughly homologous to Sld5 and Psf1, and GINS23, roughly homologous to Psf2 and Psf3) which form a 2:2 complex, while others bear a single GINS51-type protein that functions as a homotetramer (Oyama et al., 2011). *Haloferax* belongs to this second group, bearing a single GINS gene of GINS51 homology.

In Eukaryotes, the third component of the CMG complex is Cdc45, a protein that shares distant homology with bacterial RecJ. In archaea, this role is thought to be occupied by RecJ-like proteins termed GAN (GINS-associated nuclease). Like RecJ, these proteins possess a DHH phosphoesterase domain required for exonuclease activity, although this is catalytically inactive in many eukaryotic Cdc45 proteins (Srivastav et al., 2019). Studies in other archaea species seem to suggest an interaction between GAN and GINS, rather than GAN and MCM (Nagata et al., 2017; Oyama et al., 2016). In several archaeal species, GAN and GINS have been shown to form a stable complex in vitro, and addition of this complex has been shown to increase the activity of MCM in vitro (Xu et al., 2016).

Many archaea species (specifically within the euryarchaeota) possess multiple RecJ homologues, with diversified functions. This has made it challenging to identify the specific protein involved in the CMG complex in some organisms. *H. volcanii* possesses four known RecJ homologue candidates, termed *RecJ1-4*, of which 1, 3 and 4 are nonessential (Lever, 2020). RecJ3 and RecJ4 have recently been implicated in the repair of double strand DNA breaks (Jia et al., 2023), and RecJ2 seems to have diverged from the other proteins in terms of structure (Smith, 2021). Therefore, RecJ1 is currently considered the strongest candidate for the role of Cdc45-like protein in the CMG complex.

1.3.1.3 Other replisome components

Haloferax bears both bacterial-like (DnaG) and eukaryotic-like primases (composed of PriS and PriL subunits), although it is thought that only the eukaryotic-like primases are needed for DNA replication. Of these, PriS contains the active site for polymerase function, while PriL aids in initiation of primer synthesis, but is not necessary for elongation (Greci et al., 2022). *DnaG* can be deleted without affecting cell viability, although the *PriS* and *PriL* genes are essential (Le Breton et al., 2007). In *S. solfataricus*, the PriSL heterodimer has been shown to interact with GINS, suggesting that these proteins may mediate the recruitment and activity of primases in the replication fork (Marinsek et al., 2006). It may be assumed that, like other archaeal primases, *Haloferax*'s primases are able to polymerise both DNA and RNA, which may aid in handoff from the primase to the replicative DNA polymerases (Greci et al., 2022).

DNA polymerases are found in all life forms (as well as some viruses) and exhibit some diversity of form and function. These are arranged into several families. Bacteria are mainly dependent on family A and C polymerases, while the main replicative polymerases in the Eukarya are from family B (Johansson & Dixon, 2013). Archaea encode a range of DNA polymerases, including the archaea-specific family D polymerases (found in all archaeal taxa except the Crenarchaeota). The essentiality of PolB and PolD varies by species, and most species encode multiple variants of these genes (Makarova et al., 2014). It can therefore be difficult to predict which specific DNA polymerases are preferentially active at the leading and/or lagging strands during replication in a given species.

Archaeal PolBs are present in all archaea taxa, and share homology with the catalytic subunit of eukaryotic family B DNA polymerase. PolB typically demonstrates classic DNA polymerase structure, (including the palm, fingers and thumb domains), and can both polymerise and proofread DNA during replication (Makarova et al., 2014). Unlike their eukaryotic equivalent, many archaeal PolBs also include an N-terminal exonuclease domain, as well as a uracil recognition domain, which scans ahead to detect misincorporated uracil or hypoxanthine, and pauses replication four nucleotides before these sites (Killelea et al., 2010).

In several archaea species, PolB can be successfully deleted, suggesting that it is not the canonical DNA polymerase used in replication, and that this role is fulfilled by PolD DNA polymerases (Cubonová et al., 2013). However, the Crenarchaea possess no PolD, meaning that canonical replication of DNA must be performed by their several PolB enzymes. For example, in the crenarchaeon *Sulfolobus solfataricus*, it has been suggested that the PolB1 DNA polymerase synthesises the leading strand, while PolB3 synthesises the lagging strand (Bauer et al., 2012).

Archaeal PolD is made up of two subunits encoded by separate genes; DP1 and DP2. DP2 carries out the catalytic activity of DNA polymerisation, while the smaller DP1 functions as proofreader, and possesses 3'-5' exonuclease activity. In *Thermococcus kodakarensis*, both are required for maximal DNA polymerase activities (Takashima et al., 2019). It is thought that PolD may function as a lagging strand polymerase, as it preferentially binds and extends RNA-primed DNA, with much greater efficiency than PolB. It may then hand-off polymerase activity to PolB, as studies in *Pyrococcus abyssi* have shown that PolB can displace PolD from the DNA and continue synthesis (Greenough et al., 2015).

H. volcanii carries single copies of the two PolD subunit genes (PolD1 and PolD2), and two PolB genes, of which one (PolB1) is a family B3 DNA polymerase encoded on the main chromosome, and one (PolB2) is carried on pHV4. The PolB2 gene belongs to the PolB2 family, which is often inactive in archaea (Rogozin et al., 2008). Furthermore, the gene itself contains a higher percentage of rare codons, suggesting that it is a more recent acquisition; probably the result of horizontal gene transfer (Smith, 2021). This gene is non-essential, suggesting that it does not play an important role in DNA replication; however, both PolB1 and PolD have been found to be essential in *H. volcanii* (Smith, 2021). As an aside, the closely related *Haloferax mediterranei* does not possess a PolB2 gene, further supporting the more recent acquisition of this gene (Makarova et al., 2014).

The activity of DNA polymerases is greatly increased by the presence of a sliding clamp protein, which helps anchor the enzyme to the DNA strand. In bacteria, this role is played by β-clamp protein, while in the Archaea and Eukarya, the trimeric PCNA (proliferating cell nuclear antigen) is used. *H. volcanii* encodes a single gene for PCNA; its proteins interact in a head-to-tail manner to form the stable ring structure of the sliding clamp (Winter et al., 2009). PCNA is predicted to interact with a wide range of proteins possessing PIP boxes (PCNA-interacting peptide boxes), including DNA polymerases B1 and D2, Fen1 endonuclease and DNA ligase (Mayanagi et al., 2018). As PCNA is present as a trimer, it has been suggested that it interacts with multiple proteins simultaneously, forming a scaffold for other proteins and orchestrating their activity at the replication fork.

PCNA forms stable rings in solution, and therefore cannot easily be loaded onto the DNA strand. It is loaded onto the primer-template junction by RFC (replication factor C). *H. volcanii* encodes three homologues of this gene, all of which are essential (Giroux & MacNeill, 2015; Pérez-Arnaiz et al., 2020). While only a single copy of PCNA need be recruited to the leading strand, the lagging strand must be repeatedly re-loaded with PCNA by RFC as it passes Okazaki fragments. PCNA also plays an important role in the maturation of Okazaki fragments, stimulating Fen1 to cleave displaced RNA

primers, and recruiting and stimulating DNA ligase to seal the nicked duplex that results (Mayanagi et al., 2018).

Exposed single-stranded DNA at the replication fork is protected by single-stranded DNA binding proteins (SSBs). These bind on to exposed ssDNA, protecting it from chemical damage and preventing formation of secondary structures such as hairpins. *H. volcanii* encodes three homologues of a eukaryotic-type SSB termed RPA. Of these, only RPA2 is essential to survival, while RPA1 and 3 are both inessential (Skowyra & MacNeill, 2012). In the euryarchaeota, the *rpa1* and *rpa3* genes are each part of an operon with an additional protein, termed *rpap1* and *rpap3* respectively (for RPA-associated protein). Histidine-tagged RPA1 pulls down RPAP1 during protein purification, and the same is true for RPA3 and RPAP3. Deletion of *rpa3* or *rpap3* result in increased sensitivity to DNA damaging agents, including UV irradiation and MMC; no such sensitivity is observed in *rpa1* or *rpap1*-deleted strains (Stroud et al., 2012). Equivalent sensitivity is observed in both *Δrpa3* and *Δrpap3* strains, suggesting that the two proteins are part of the same pathway. The exact role of RPAP is yet to be fully understood.

A summary of the *Haloferax* replication fork (not including Orc proteins or origin) is shown in Figure 3, below.

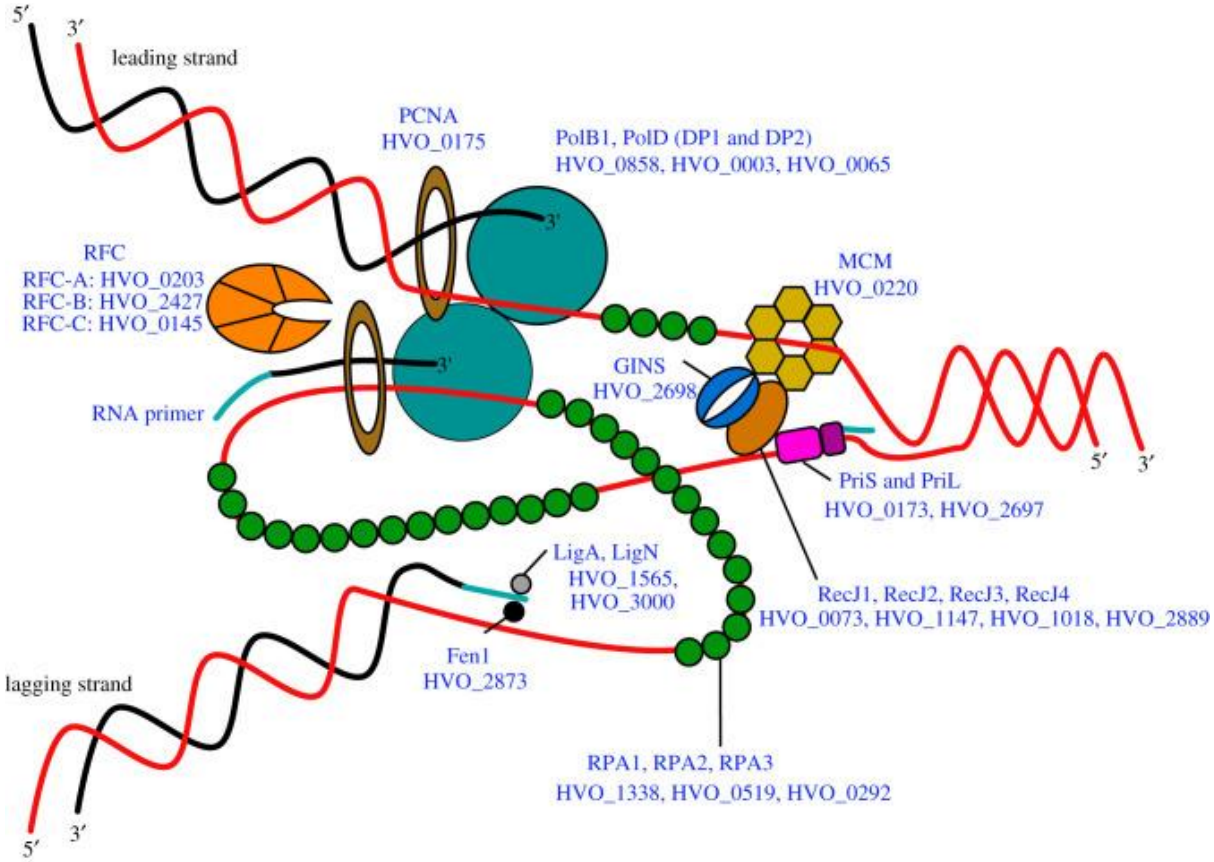


Figure 3. Overview of the Haloferax *replication fork, taken from Pérez-Arnaiz et al. (2020).*

1.4 DNA repair

1.4.1 DNA damage

DNA damage is commonplace and largely unavoidable. There are many factors that can contribute to DNA damage; some environmental (exogenous) but many spontaneous (endogenous). While some low levels of unrepaired DNA damage may be tolerated, high levels of unrepaired DNA damage can rapidly lead to mutagenesis, genome instability and eventually death (Hakem, 2008).

Normal metabolic activities within the cell can result in a variety of forms of DNA damage. Free radicals produced as a result of respiration are highly reactive, and can cause around 100 types of lesions to bases and the sugar-phosphate backbone (Chatterjee & Walker, 2017). One of the most studied of these lesions is conversion of guanine to 7,8 dihydro-8-oxoguanine (8-oxo-G) by hydroxylation. 8-oxo-G pairs with adenine instead of cytosine, resulting in mutagenesis if not repaired before DNA replication.

Other types of spontaneous DNA damage include deamination of bases, which can result in conversion of cytosine, adenine, guanine and 5-methylcytosine to uracil, hypoxanthine, xanthine and thymine, respectively. While this is a spontaneous decay - the result of chemical processes or enzymatic activity - deamination occurs at a much higher frequency in single-stranded than double-stranded DNA (Lindahl, 1993). In addition to spontaneous deamination, abasic sites can be formed by hydrolysis of the bond connecting the nitrogenous base to the sugar backbone of the nucleotide (Chatterjee & Walker, 2017). Errors can also be introduced by DNA polymerases during DNA replication, resulting in mismatches in the duplex.

Exogenous sources of DNA damage can be physical or chemical in nature. UV irradiation can cause pyrimidine dimers; bulky lesions that distort the duplex and can block the activity of DNA and RNA polymerases (White & Allers, 2018). Chemicals that can damage DNA include a range of agents that can modify the DNA. Notable members include intercalators and inter- and intrastrand crosslinkers, as well as alkylating and methylating agents (Chatterjee & Walker, 2017).

The double strand break is one of the most damaging forms of DNA damage. In addition to the obvious barriers they pose to cellular processes (such as transcription and replication), incorrect repair of these structures can result in large-scale chromosomal rearrangements and mutagenesis (Cannan & Pederson, 2016).

The most common varieties of DNA damage, as well as the mechanisms by which they are typically repaired, are summarised in Figure 4, below.

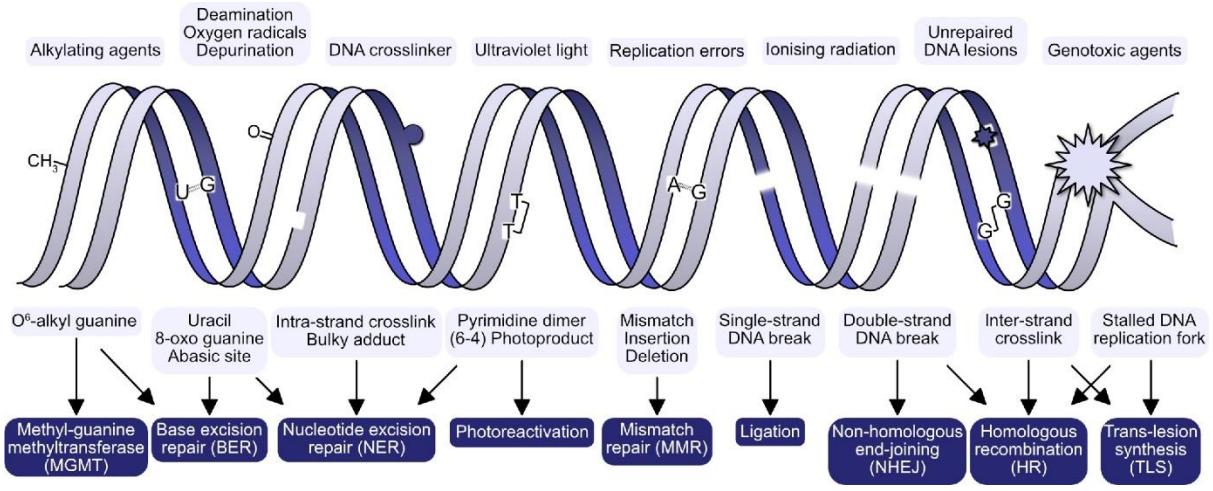


Figure 4. Common types of DNA damage and their typical repair pathways. From White & Allers (2018).

The rate of mutation in *H. volcanii* has been estimated to be around 3.15×10^{-10} per site per generation, or around 0.0012 per genome per generation (Kucukyildirim et al., 2020). This implies a robust DNA repair machinery, especially given an environment that would be challenging for many other organisms.

1.4.2 DNA repair pathways

Archaea possess many pathways for DNA repair; some are similar to those used by eukaryotes, while others may have been acquired from bacteria through horizontal gene transfer (White & Allers, 2018).

1.4.2.1 Direct repair

In many cases, repair can be effected through direct reversal of the chemical damage. For example, UV-induced dimers and photoproducts can be repaired through photoreactivation, in which photolyases can utilise energy from blue or violet light to excite cofactors. The cofactors can then donate an electron to the dimer, generating a radical anion which causes the two bases to split (Zhong, 2015). This pathway is likely important for organisms such as *H. volcanii*, whose saltern habitats typically involve little shelter from UV exposure. Many other forms of DNA repair can be performed by specific enzymes which reverse the chemical alterations the DNA has suffered.

Many lesions can be repaired either through base excision repair (BER) or nucleotide excision repair (NER). In these pathways, excision of the damage is performed; removing either the individual damaged base (removed by a specific glycosylase) or a short stretch of bases from the affected strand (Grasso & Tell, 2014). Removal of the damaged base (BER) is followed by cleavage of the backbone, allowing access for a DNA polymerase to synthesise either a new base, or a short stretch of bases (White & Allers, 2018). The backbone is then re-sealed by DNA ligases, which in archaea can be ATP-dependent, like those found in eukarya, or NAD+-dependent, like those found in bacteria (Martin & MacNeill, 2002; Zhao et al., 2006). The small gaps in the DNA produced during BER can be filled by polymerase X. This polymerase is not essential in *H. volcanii* (T. Allers, unpublished data).

NER follows a similar pathway to BER, but involves removal of a longer stretch of nucleotides from the side of the damaged bases. In *H. volcanii*, this is carried out by the UvrABC proteins, similar to those found in bacteria (White & Allers, 2018). However, *H. volcanii* also encodes some proteins that are homologous to NER proteins seen in eukaryotes; specifically, the Hef endonuclease (homologous to eukaryotic XPF) and Fen1 (homologous to eukaryotic XPG) (Lestini et al., 2010). Unlike its

eukaryotic cousin, Hef is not thought to play a role in NER, having instead been implicated in rescue of stalled replication forks.

1.4.2.2 Double-strand break repair

As mentioned above, double-stranded breaks (DSB) represent a severe form of DNA damage that can interfere with many processes. Several methods of repair are available in cases of DSB, of which the most accurate is homologous recombination (HR). However, HR is an involved and energetically complex process, so a mixture of repair pathways are likely within the cell (Pérez-Arnaiz et al., 2020).

In both bacteria and eukaryotes, the most direct pathway to resolve a DSB is non-homologous end joining (NHEJ) (Lieber, 2010). In this pathway, Ku proteins recognise the break, and recruit a selection of other components to join the two broken ends together. In both bacteria and eukaryotes, this pathway is both convenient and necessary prior to S-phase of the cell cycle, when only one genome copy is present. Under such circumstances, there is not necessarily a homologous chromosome to act as a template, and indeed, there is a finite number of DNA molecules that could be the source of the broken ends. While the mechanism of joining may introduce new mutations, the chromosome will likely be restored.

Most archaeal genomes identified do not contain Ku proteins (Blackwood et al., 2013). No Ku protein is present in *H. volcanii*, and this organism is not thought to perform NHEJ (Pérez-Arnaiz et al., 2020). While it may appear unusual, this makes sense in light of *H. volcanii*'s genome structure; the presence of many copies of the genome could result in multiple DSB present in the cell at once. Joining all of these broken ends without any concern for their homology could rapidly cause large-scale chromosomal changes and genome instability. *H. volcanii* can instead repair DSB through microhomology-mediated end joining (MMEJ).

In MMEJ, the ends of DSB are resected by Rad50 and Mre11 to produce ssDNA overhangs. Small regions of homology between these overhangs can result in annealing of the two strands. The duplex can then be restored through a combination of cleaving overlaps, gap filling, and ligation (Delmas et al., 2009). A diagram of this process, and the identified enzymes involved in *H. volcanii*, are shown in Figure 5, below.

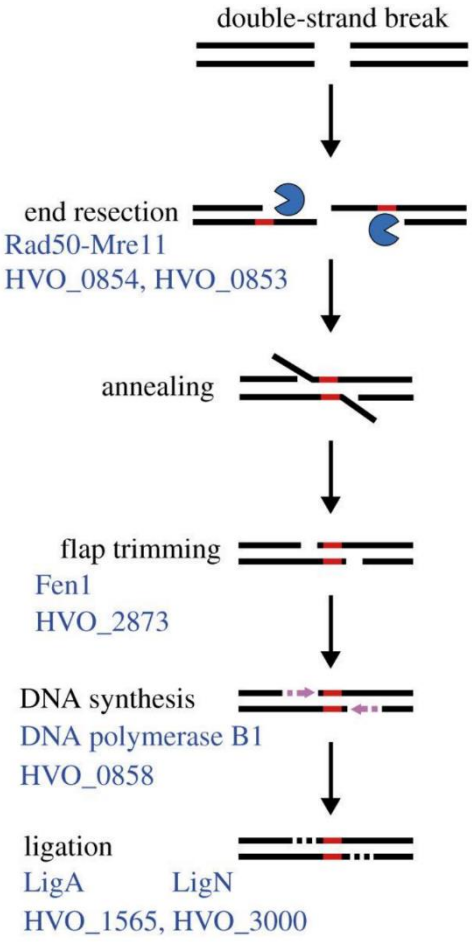


Figure 5. Repair of DSB by MMEJ. Names of proteins identified in this process in H. volcanii *are shown in blue. Adapted from Pérez-Arnaiz et al. (2020)*

While MMEJ is a relatively fast and simple method for repair of DSB, it has a tendency to introduce small deletions and other errors. The gold standard for accurate repair of DSB is homologous recombination, which is more time consuming and costly for the cell. It has been suggested that MMEJ is frequently used in *H. volcanii* as a "stop-gap" mechanism; rapidly repairing damaged strands, while HR can take place later to revert any introduced mutations (Delmas et al., 2009).

1.4.2.3 Homologous recombination (HR)

HR is commonplace in all domains of life, and is the most accurate pathway for repair of DSB. In simple terms, HR is a process to repair DNA damage using a template DNA strand to ensure accuracy. This process falls into three steps; processing of the broken ends to generate ssDNA overhangs (pre-synapsis), invasion of a homologous sequence to form a displacement loop (D-loop; synapsis) and finally, strand synthesis and separation of the two chromosomes (post-synapsis) (Pérez-Arnaiz et al., 2020).

The first step of pre-synapsis is recognition of the DSB by the Mre11-Rad50 complex. These resect the end of the break, forming a small ssDNA overhang. Both *rad50* and *mre11* are inessential in *H. volcanii*, and interestingly, their deletion improves resistance to several types of DNA damage (Delmas et al., 2009). This may be due to a lack of these proteins increasing the rate of MMEJ, as opposed to HR, which is a faster and more cost-effective method to repair damage. In a highly polyploid organism like *H. volcanii,* widespread HR could result in complex interconnections of many chromosomes, and could cause complications during cell division.

ssDNA overhangs at the DSB ends are then coated in RadA, a recombinase protein homologous to eukaryotic Rad51 and bacteria RecA, although bearing more similarity to Rad51 (Lin et al., 2006). *H. volcanii* also carries a paralogue of this protein, *radB*, which is found only in the Euryarchaea. Deletion of either gene from *H. volcanii* confers a significant growth defect, suggesting an important role for these proteins (and HR in general) in this species (Wardell et al., 2017). RadB aids in loading and polymerisation of RadA onto the ssDNA to form a nucleoprotein filament. This filament is then able to seek out regions of homology in other DNA duplexes. Where these are found, RadA mediates strand invasion, displacing one strand of the duplex and forming a D-loop. This is shown in Figure 6, below.

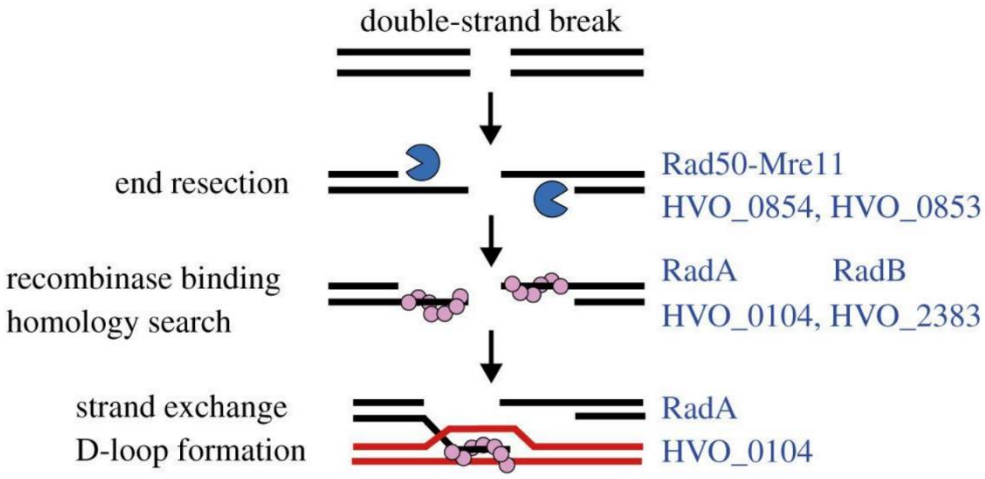


Figure 6. Double strand break repair by homologous recombination. Homologous chromosome is shown in red. Names of proteins implicated in each step in H. volcanii *are listed in blue. Adapted from Pérez-Arnaiz et al. (2020).*

Following D-loop formation, the end of the DSB can act as a primer for the further synthesis of this strand, using the homologous chromosome as a template. Studies in *Pyrococcus abyssi* have shown that, in this species at least, both PolB and PolD are capable of extending strands following RadA-mediated strand invasion (Hogrel et al., 2020). As synthesis progresses, it can further displace the other strand of the duplex (where present), extending the D-loop, and potentially allowing the displaced strand to act as template to the partner DSB.

In archaea, the Hel308 helicase has been implicated in homologous recombination (discussed in more detail in section 1.5, below). This helicase can load onto ssDNA, and can displace the complementary strand or bound proteins during its ATP-powered translocation in a 3'-5' direction (Guy & Bolt, 2005; Richards et al., 2008). Despite some homologues of this protein bearing the name Hjm (for Holliday junction migration), *in vitro* experiments have suggested a preference for unwinding forked DNA structures and D-loops over Holliday junctions (Fujikane et al., 2006; Guy & Bolt, 2005). However, pulldown analyses in *Sulfolobus tokodaii* have revealed an association between this protein and the Holliday junction resolvase Hjc (Li et al., 2008). Hel308 may play a role in progress or resolution of HR, potentially unwinding D-loops and allowing re-formation of the original duplex. This allows the two ends of the DSB to then anneal and be repaired by DNA synthesis, using the newly-synthesised strand as a template. Ligases can the seal the backbone of the repaired duplex. This outcome is shown in Figure 7, below.

During HR, four-armed DNA structures known as Holliday junctions can be formed, and may migrate in position relative to the sequences on the strands involved. These must eventually be cleaved to separate the two homologous chromosomes. Resolvases are found in all three domains, but in archaea this role is thought to be played by Hjc (for Holliday junction cleavage) (Lestini et al., 2010).

Holliday junctions cleave two opposing strands simultaneously to free the two duplexes from each other; depending on the direction of the cut, this may generate crossover or non-crossover products. The gene encoding Hjc is not essential in *H. volcanii,* suggesting that this species possesses an additional method for resolving Holliday junctions (Lestini et al., 2010).

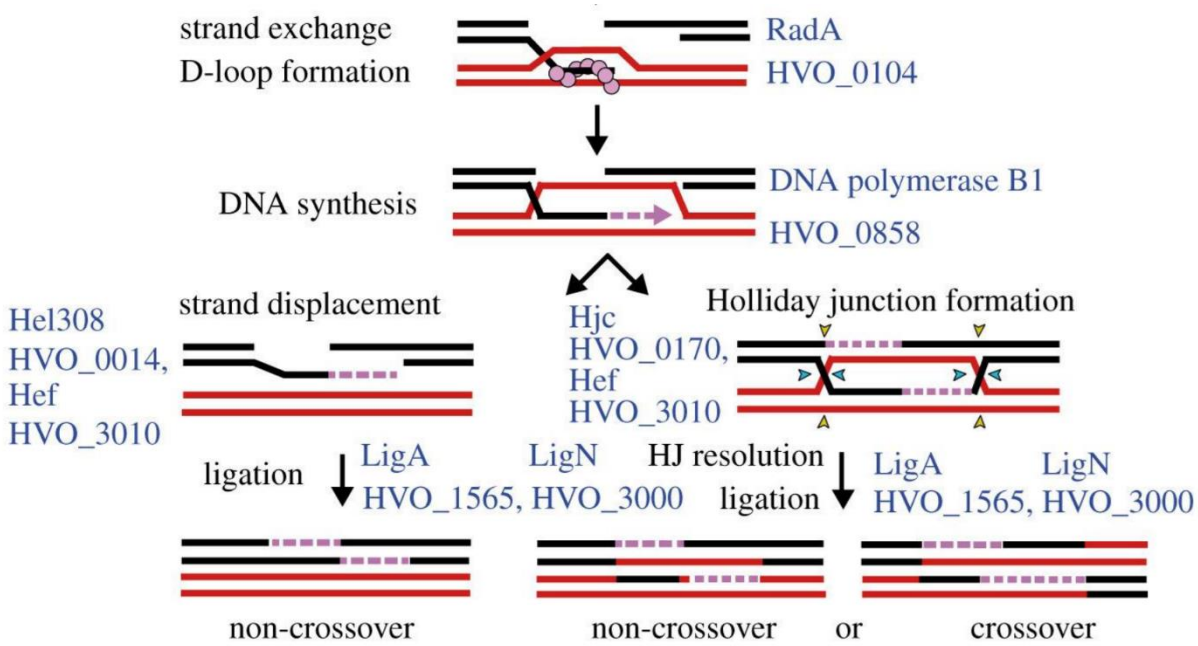


Figure 7. Potential outcomes of HR. Homologous chromosome is shown in red. Names of proteins implicated in the relevant steps in H. volcanii *are shown in blue. Adapted from Pérez-Arnaiz et al. (2020)*

Hef (helicase-associated endonuclease fork-structure DNA) is a euryarchaeal-specific endonuclease (homologous to eukaryotic XPF) (Komori et al., 2004). While XPF is implicated in NER, Hef has been shown to cleave a range of forked or branched DNA structures, several of which are present in Holliday junctions and stalled forks (discussed in section 1.4.3, below) (Lestini et al., 2013). It is thought that this may represent a parallel, independent pathway for resolution of Holliday junctions; both Hjc and Hef are inessential in *H. volcanii*, but deletion of both cannot be tolerated (Lestini et al., 2010). In the euryarchaea, Hef is thought to function alongside HAN (Hef-associated nuclease), which co-ordinates or mediates its function (Nagata et al., 2017).

Following the resolution of the Holliday junctions, ligases can seal the resultant nicks in the DNA backbones, resulting in the restoration of both homologous chromosomes, with the double-strand break repaired. Note that, if the two chromosomes vary at some of the alleles used as template for DNA synthesis, this pathway can result in gene conversion or exchange of alleles, in the case of crossover outcomes.

1.4.3 Stalled fork repair

Many factors can obstruct the progression of replication forks; for example, DNA lesions, inter-strand crosslinks, nicks to the backbone, or collision with other DNA-bound proteins can all result in the replication fork stalling or collapsing. While translesion synthesis may occur, allowing bypass of the problematic sequence, some instances of stalled DNA polymerases may need greater attention in order to be resolved (Kondratick et al., 2021).

Stalled forks cannot be allowed to remain, as they prove a barrier to further replication, transcription, and segregation of chromosomes during cell division. Depending on the cause of the event, they may also involve long ssDNA stretches, which are vulnerable to further damage (Pacek et al., 2006), or may involve broken DNA strands. The mechanisms available to remedy stalled forks are several, and are summarised briefly in Figure 8 below, from Lever (2020).

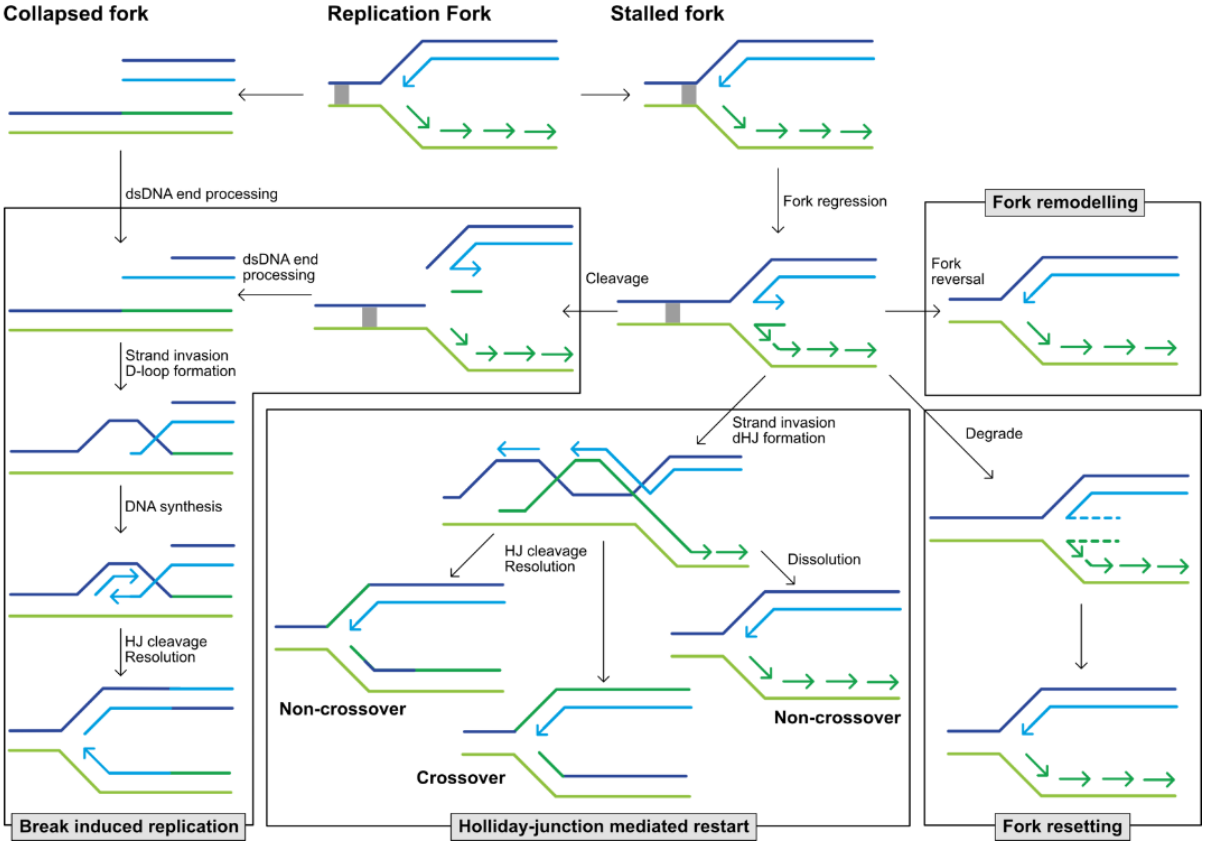


Figure 8 Options for restart of replication forks. A barrier to replication is indicated by the grey block. HJ: Holliday junction. dHJ: double Holliday junction. From Lever (2020).

1.4.3.1 Non-recombination processing of stalled forks

Stalled forks can be regressed into a four-way Holliday junction intermediate known as a "chicken foot". This is achieved by complementary pairing of the nascent leading strand with the nascent lagging strand to form a fourth duplex opposite the direction of fork progression. In bacteria, this regression of the fork is thought to be initiated by RecG or RecA (McGlynn et al., 2001; Robu et al., 2001), while in humans the RecQ family helicases BLM and WRN are thought to be responsible (reviewed in Liao et al., 2018). The picture of this mechanism in archaea is currently incomplete, but the Hel308 homologue from *Sulfolobus tokodaii* (Hjm) is capable of regressing stalled forks *in vitro* (Li et al., 2008). However, in some cases Hjm may preferentially unwind the lagging strand of the stalled fork instead (again, *in vitro*) (Guy & Bolt, 2005). Hef (helicase-associated endonuclease for fork-structured DNA) from *Pyrococcus furiosus* contains both a DEAD/H helicase and an endonuclease domain, and is hypothesised to migrate and then resolve Holliday junctions (Komori et al., 2004).

In addition to providing protection and stability for single-stranded DNA, this can result in the junction "backing up" from the site of the stall, allowing opportunity for the cause of the stall (be it DNA lesion or bound protein) to be resolved. There is also the possibility of template switching,

wherein the nascent leading strand is able to use the lagging strand as a template or vice versa, potentially bypassing the site of the DNA lesion.

The fork can then be restored, either be reversal of the regression or by degradation of the regressed arms. These various rearrangements of the fork are often referred to as “remodelling”. An example is shown in Figure 9, below.

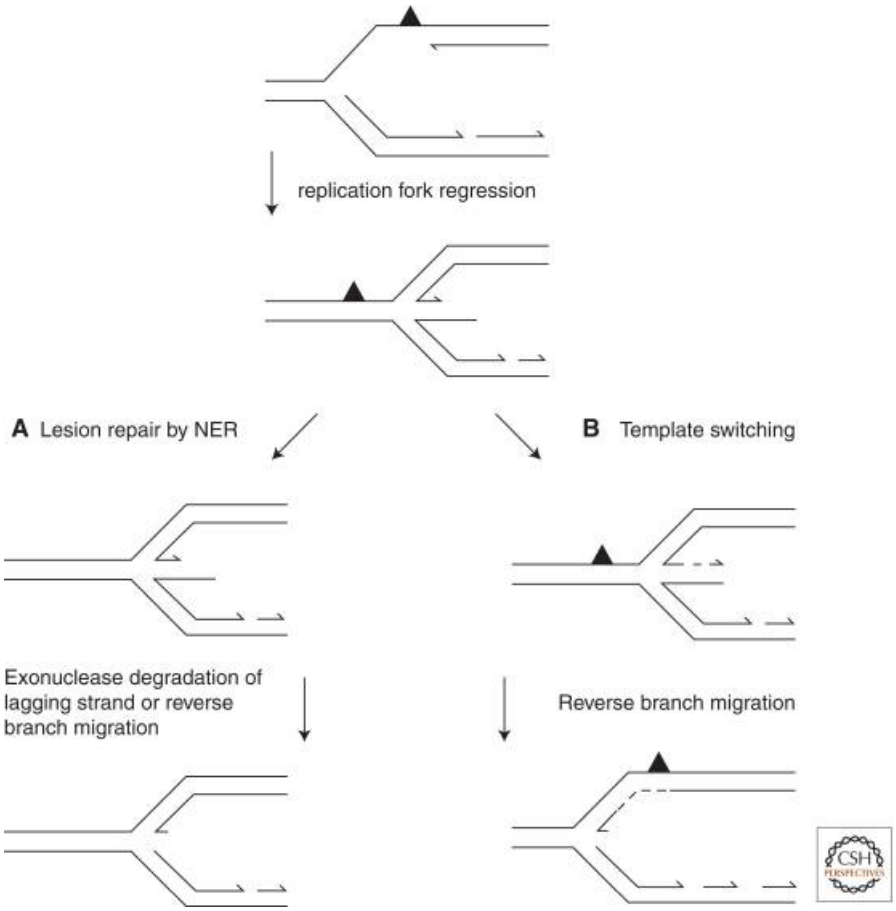


Figure 9. Model for fork regression following stall due to lesion on the leading strand. Lesion is represented by the black triangle. Adapted from Yeeles et al. (2013).

Following resolution of the lesion and fork remodelling, either by degradation of the fourth arm or by migration of the Holliday junction equivalent, the normal structure of the replication fork is restored. In bacteria, degradation of the regressed strands is carried out by RecJ nuclease, and in humans, the DNA2 nuclease/helicase (Pérez-Arnaiz et al., 2020). It is possible that, in archaea, Hef (and its interactor HAN) is able to play a similar role, as it is capable of both helicase and endonuclease abilities (Nagata et al., 2017). Despite the presence of a Holliday junction-like structure, no recombination to another chromosome has occurred as a result of fork regression, remodelling or resetting. These are the mechanisms outlined in Figure 9, above.

1.4.3.2 Processing of stalled forks by double Holliday junction

Following regression of the stalled fork, the regressed end may be processed in a similar manner to a DSB (discussed above). Resection of the end to generate a ssDNA overlap (if necessary) can allow formation of a RadA nucleoprotein filament that can seek out homology on the template strand. This may allow bypass of the original lesion or replisome block, and can also provide a template to the

other stalled strand. As with homologous recombination, described above, this may be resolved by dissociation of the strands and unwinding of the D-loops. Alternatively, if the Holliday junctions persist, they must eventually be resolved to separate the various strands. This can restore the replication fork.

1.4.3.3 Break-induced replication

If the replication fork passes an unrepaired nick in the DNA backbone, or if the chicken-foot structure caused by fork regression is cleaved by a Holliday junction resolvase, a double-strand break can occur, effectively snapping off one arm of the replication fork. While this should leave one intact duplex (albeit with a nick in the backbone), the broken arm bears a double-stranded break, but no corresponding “other end” of the break is in evidence. This is typically repaired through recombination with an intact chromosome, and is known as break-induced replication.

Firstly, the ends of the broken strand must be processed to generate a 3’ overhang of ssDNA (if one is not already present). Recombinases can then catalyse strand invasion, forming a nucleoprotein filament that searches for complementary sequences (Malkova & Ira, 2013). The high ploidy often observed among archaea is likely of benefit in this stage, offering multiple candidates for strand exchange. The invading 3’ ssDNA end can then be used as a primer for new DNA synthesis. Following strand invasion and the initiation of DNA synthesis using the invaded strand as template, several options are possible. Resolution of the Holliday junction can result in reconstitution of the replication fork, as shown in Figure 10(i), below. In other potential resolutions of break-induced repair, the invading strand is replicated in a migrating bubble (Saini et al., 2013). In yeast, break-induced replication has been observed continuing for hundreds of kilobases, suggesting that it is a relatively stable mode of replication (Malkova et al., 2005).

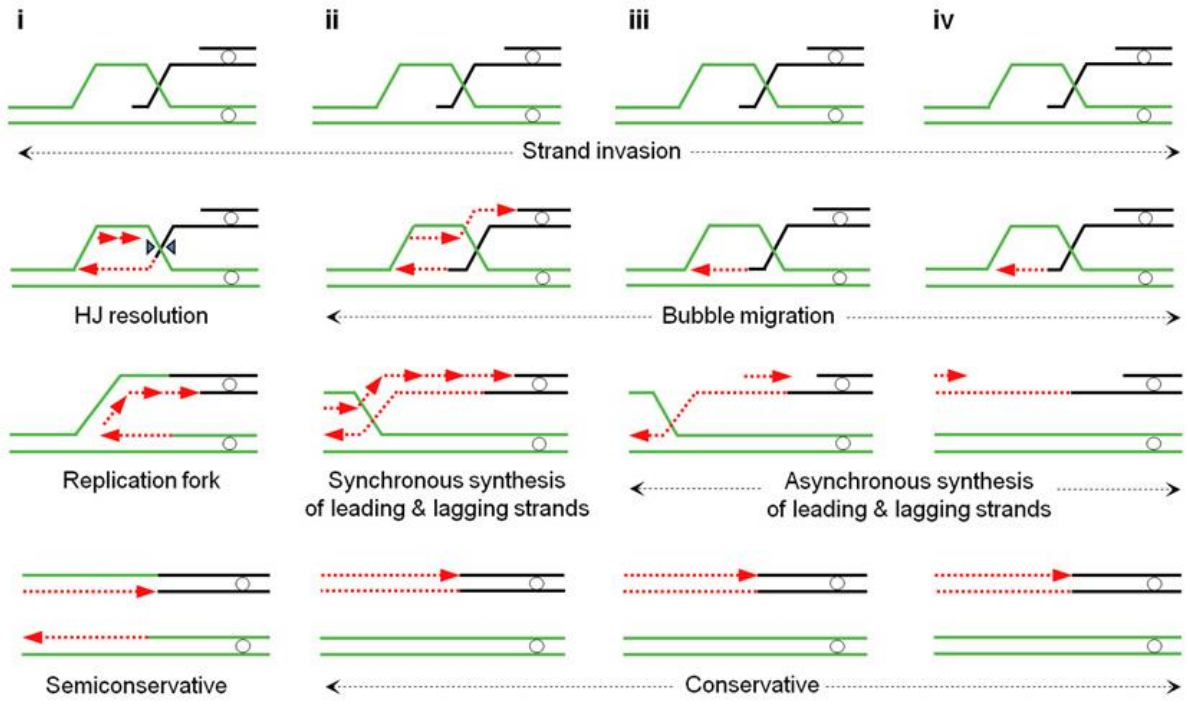


Figure 10. Potential resolutions of break-induced replication. Following end resection of the DSB, and strand invasion of the overhang, four resolutions of break-induced replication are available, labelled i-iv above. In resolution (i), resolution of the Holliday junction results in formation of a typical replication fork. In resolution (ii), the invading strand is replicated in a migrating bubble, while the lagging strand uses the newly-synthesised strand as a template. This results in roughly synchronous synthesis of the two strands. In resolution (iii) and (iv) however, the leading and lagging strands are synthesised asynchronously, with the lagging strand using the leading strand as template. This potentially results in the leading strand existing as ssDNA for some time. Adapted from Saini et al. (2013).

Notably, three of the listed outcomes result in conservative replication, wherein original parent strands remain paired after synthesis, as do the newly-synthesised strands. This is in contravention of accepted dogma of semi-conservative DNA replication, as originally posited by Meselson and Stahl in 1958 (Meselson & Stahl, 1958). In addition, the resolution shown in Figure 10(iv) involves only leading-strand synthesis.

In addition to these quirks of replication, resolutions (iii) and (iv) in Figure 10 potentially involve long stretches of time in which ssDNA is present, before the synthesis of the second strand. This is likely the source of the much higher frequency of mutation observed in yeast strains undergoing break-induced replication (Saini et al., 2013).

In both human and yeast cells, the PIF1 helicase is important in BIR initiation and long-track DNA synthesis by this mechanism (Li et al., 2021). It is unknown which enzymes fulfil this role in archaea.

1.5 Helicases

1.5.1 Overview

Helicases are a group of important enzymes involved in a range of processes within the cell. These enzymes translocate along the DNA duplex, breaking the hydrogen bonds that bind complimentary strands together, and generating single stranded DNA (or RNA). This is an ATP-dependent process, and results in unidirectional translocation along the DNA strand; helicases are therefore classified as either 5’-3’ or 3’-5’. Separation of the DNA duplex is an important part of many processes of nucleic acid metabolism, as it allows access to the nucleotides. It is therefore an integral part of DNA

replication, repair and transcription, RNA maturation and splicing, as well as processing of recombination intermediates.

Helicases are typically identified based on the presence of conserved helicase motifs, although not all proteins bearing helicase motifs are functional helicases or translocases (examples include restriction enzymes such as EcoR124I (Lapkouski et al., 2009)). Of the seven conserved helicase motifs, only the Walker A and B motifs, which are involved in ATP hydrolysis, are common to all helicases (Byrd & Raney, 2012).

1.5.2 Hel308 helicase

Archaeal Hel308 is a superfamily 2 (SF2) 3'-5' helicase. SF2 is one of the largest families of helicases, and is represented in the bacteria by RecG and homologues. SF2 helicases in metazoans include the human RecQ family helicases (including Blooms Syndrome helicase BLM and Werner Syndrome helicase WRN), mutations in which can cause genetic instability though dysregulation of recombination (Croteau et al., 2014). The human *hel308* homologue is Helicase PolQ-like, better known as HELQ (although occasionally referred to as Hel308 or POLQ-like). This helicase plays important roles in DNA repair; particularly with regard to DNA crosslinking damage (Anand et al., 2022). It has also been shown to colocalise to stalled forks alongside proteins involved in homologous recombination; Rad51 and FANCD2 (Tafel et al., 2011). The many roles of this human protein are reviewed in Tang et al. (2023). Hel308 also shares homology with the helicase subunit of the human PolΘ (encoded by the *polQ* gene), and *Drosophilia* Mus308, two unusual enzymes that include both DNA polymerase and helicase activities in different domains of the same protein (Oyama et al., 2009).

The Hel308 family is conserved throughout both metazoans and archaea, but is not found in bacteria or fungi (Byrd & Raney, 2012). In some archaea, Hel308 is better known as Hjm, for Holliday junction migration (Fujikane et al., 2005). However, it has been shown *in vitro* to exhibit a preference for forked structures over Holliday junctions (Guy & Bolt, 2005). It should be noted that in *Sulfolobus tokoddaii,* Hjm has been observed unwinding DNA in both 3'-to-5' and 5'-to-3' directions, which is unusual among this family (Li et al., 2008).

No crystal structures are available for Hel308 in *Haloferax* species, but some insight into its structure and behaviour can be inferred from sequence analysis and homologues in other Archaea. At time of writing, crystal structures of homologous proteins from three archaeal species are available; Hel308 from the euryarchaeon *Archaeoglobus fulgidus*, both apo and DNA-bound (Büttner et al., 2007), Hel308 from the crenarchaeon *Sulfolobus solfataricus* (Richards et al., 2008), and Hjm from the hyperthermophilic euryarchaeon *Pyrococcus furiosus*, both apo and with ATP or synthetic ATP homologues (Oyama et al., 2009). Although the three homologues in question only share around 30% residue identity, they demonstrate a highly similar conserved structure, shown in Figure 11, below (Oyama et al., 2009).

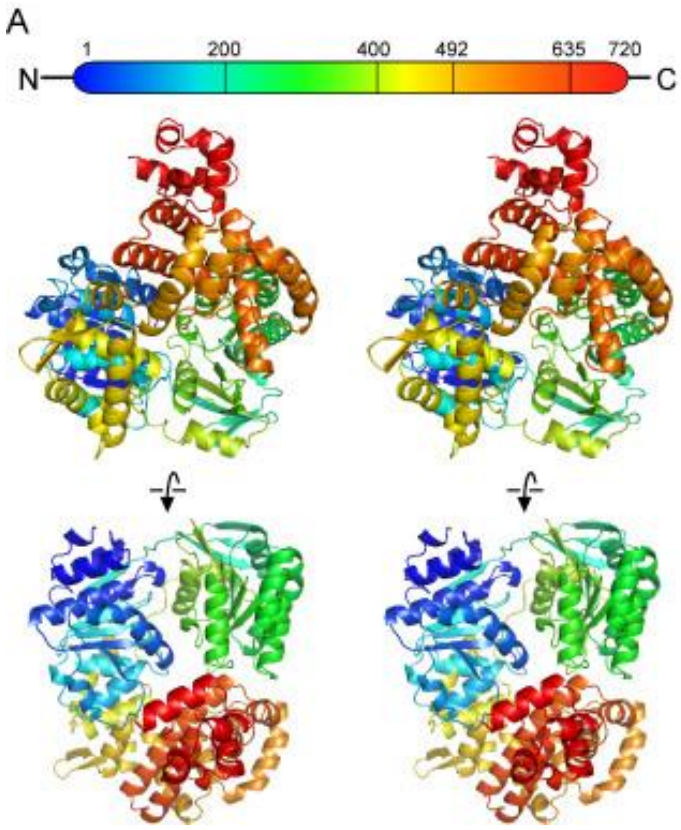


Figure 11. Structure of Pyrococcus furiosus *Hjm. From Oyama et al. (2009).*

The available Hel308 crystal structures describe a protein possessing five domains. Four of these domains (1-4) form a rough ring around a central pore, through which a single DNA strand may be threaded. This ring is not covalently closed, allowing loading of the enzyme onto the DNA strand (Richards et al., 2008). Domain five sits slightly apart from the core, in a position to make contact with ssDNA exiting the pore. Each of these domains are discussed below.

Domains 1 and 2 (residues 1-197 and 200-416 in *H. volcanii*) are easily identifiable as ATP-binding motor domains; specifically RecA-fold domains (Woodman & Bolt, 2009). The motifs that characterise superfamily 1 and 2 helicases are found in these domains, which form the ATP binding pocket. Hydrolysis of ATP is thought to induce conformational changes leading to translocation and therefore duplex unwinding (Richards et al., 2008). Conformational changes in domains one and two are thought to bring about corresponding changes in the ratchet domain through a point of contact made with domain 2's conserved motif IVa. A study using *Methanothermobacter thermautotrophicus* Hel308 *in vitro* found that mutation of this motif resulted in increased DNA binding affinity and helicase activity (Lever et al., 2023).

In the Hel308-DNA co-crystal structure shown in Figure 12 (Büttner et al., 2007) the two RecA domains (shown in yellow and green) can be seen making contact with the single-stranded DNA as the single strand is fed through the ring structure. Signature motifs are indicated in the schematic in Figure 12a. The displaced strand appears to be guided around the exterior of the protein. The β-hairpin motif of domain 2 can be seen lodged between the parting strands of the duplex, and appears to be instrumental in guiding the two strands to separate, as seen in Figure 12b and c. The β-hairpin is also present in distantly related helicases such as NS3 of the hepatitis C virus, but absent from human HelQ (Büttner et al., 2007).

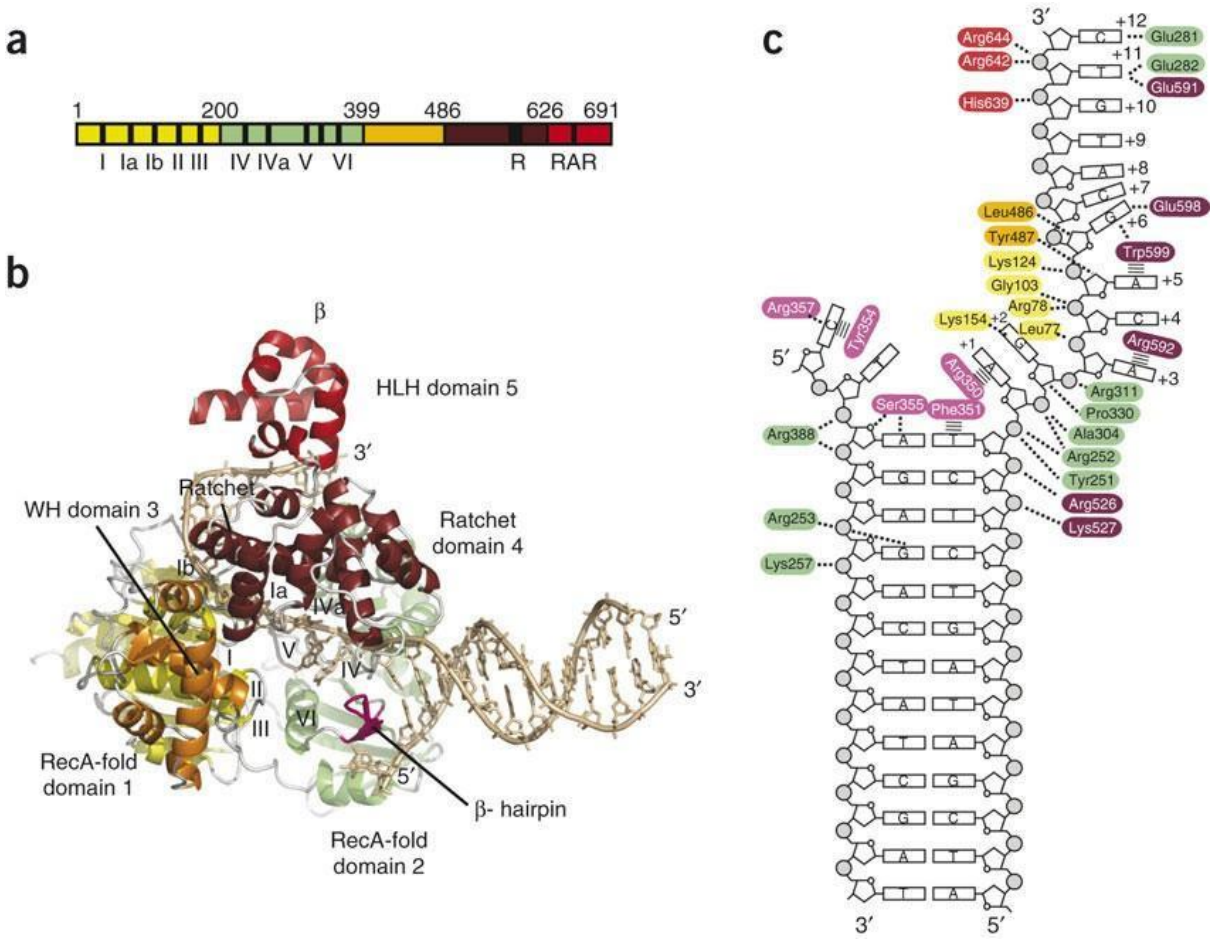


Figure 12. (a) Schematic of the five domains of Hel308. Domain boundaries are indicated on top, sequence motifs beneath. Roman numerals, SF2 helicase motifs; β, β-hairpin loop; R, ratchet helix; RAR, Arg-Ala-Arg motif. (b) Hel308 (ribbon, coloured as in a) in complex with the 15-base-pair DNA duplex and 10-base single-stranded 3' tail (beige sticks). (c) Schematic showing key interactions (dashed lines) of Hel308 with the partially unwound DNA substrate. Residues are coloured by domain (magenta, β-hairpin motif). Stripes show stacking of side chains on bases. (Buttner et al, 2007).

The Hel308-DNA co-crystal structure shown above was created by incubation of the two components in the absence of ATP. However, separation of two base-pairs of the duplex can be observed, suggesting that the melting of the duplex is not dependent on ATP. The β-hairpin loop is positioned at the boundary between the still-duplexed DNA and the melted strands, which bind to residues of the protein rather than their opposite strand. Büttner et al., (2007) posit that the energy involved in separating the duplex is sourced from binding of the nucleic acids to the enzyme, rather than from the binding of ATP, as has been suggested in reference to some other helicases, such as NS3 (Levin et al., 2005).

The above figure also demonstrates that Hel308 is capable of loading onto the ssDNA in the absence of ATP or assistance from other proteins. The 3’ ssDNA overhang can be seen passing through the pore in the centre of the protein; being a 3’-5’ helicase, the enzyme moves itself along this strand, parting the duplex where present. As translocation is assumed to require ATP, it seems likely that the helicase loads onto the ssDNA at the border between the ssDNA and the duplex, rather than the ssDNA end “threading” through the protein. This is consistent with studies of the human homologue HelQ, which appears to load itself onto the 3’ ssDNA adjacent to the duplexed DNA (Jenkins et al., 2021). Richards et al. (2008) speculate that flexibility between domains 2 and 4 is the mechanism by which Hel308 loads onto the ssDNA.

Domain 3 (residues 426-501 in *H. volcanii*) has been identified as a divergent winged helix (WH) domain (Woodman & Bolt, 2011). The role of this domain in Hel308 is not yet fully understood, being somewhat divergent from conserved amino acid sequences, and occupying a slightly unusual position within the protein relative to both the other domains and the DNA substrate (Woodman & Bolt, 2011). WH domains are conserved through a variety of helicases, and are considered a more versatile variant of the DNA-binding helix-turn-helix motif. These domains are often associated with nucleic acid binding; Richards et al. (2008) liken domain 3 to a histone protein. These structures are sometimes capable of identifying specific DNA sequences (as seen in forkhead family transcription factors (Brent et al., 2008)).

Winged helices can also be involved in a wide range of functions and protein-protein interactions (reviewed in Harami et al., (2013)). In *Methanothermobacter thermautotrophicus*, conserved acidic residues in domain 3 have been implicated in protein-protein interactions between Hel308 and the single-stranded DNA binding protein replication protein A (RPA) (Woodman et al., 2011). This may aid in recruitment or loading of Hel308 onto ssDNA.

Domain 4 (502-644) is believed to be the ratchet domain. It is composed of a 7-helical bundle of uncertain homology. The structure is quite rare, but it is posited by Richards et al. (2008) that it may function as an ATP-powered ratchet in unwinding the DNA. Mutational studies using *Methanothermobacter thermautotrophicus* Hel308 have suggested that close contact between the ratchet domain and motif IVa of domain 2 couples ATPase-induced conformational changes of the RecA domains into movement of the ratchet (Lever et al., 2023). Mutation of the proposed point-of-contact residue (Y586 in *M. thermoautotrophicus*) resulted in a hyper-helicase protein, but did not affect ability to anneal complementary ssDNA strands together.

Domain 5 (647-705) contains a helix-hairpin-helix motif, a structure often associated with DNA binding. This domain is clearly visible as sitting slightly apart from the main body of the protein. The highly conserved RxRAR motif resides in this domain, and the central of the three arginine residues can be seen interacting with the emerging ssDNA tail in the *A. fulgidus* co-crystal structure. Büttner et al. (2007) suggest that interactions between this domain and the complementary strand of the duplex may aid in identification of forked DNA structures. The first and third arginine residues in this motif (underlined RxRAR) are thought to make functionally important contacts with the ratchet domain; disruption of these residues leave the enzyme unable to function as a helicase (Woodman et al., 2011). Richards et al. (2008) posited that domain V is an autoinhibitory domain or molecular brake, as truncated proteins lacking this part appear to exhibit much faster helicase activity than the full-length protein.

1.5.2.1 Hel308 biochemistry

Hel308 (or Hjm) was first discovered in *P. furiosus* through its ability to complement *E. coli recQ* mutants (Fujikane et al., 2005). Further study revealed a preference for binding DNA structures with single-stranded components, particularly forked structures. When interacting with a forked DNA structure, *P. furiosus* Hel308 exhibits a preference for unwinding the lagging strand equivalent (Fujikane et al., 2006). The *M. thermautotrophicus* homologue is thought to behave similarly, and is thought to be unable to load onto intact duplex DNA, only ssDNA (Guy & Bolt, 2005). Studies in *Sulfolobus* solfataricus have also shown that, while translocating along DNA, Hel308 can displace proteins from the DNA strand (Richards et al., 2008). Hel308 may therefore be well placed to interact with stalled replication forks, remodelling the fork and perhaps dislodging proteins present, although its precise interactome is not yet fully characterised.

Sulfolobus tokodaii Hjm has been confirmed to interact with the Holliday-junction specific endonuclease Hjc by pulldown and yeast two-hybrid assay (Li et al., 2008). Hjc was also shown to inhibit Hjm helicase activity. Hjc is an archaeal endonuclease specific to four-way junctions, and is thought to act as a dimer to symmetrically cleave and thus resolve Holliday junctions (Komori et al., 2000). This firmly places Hel308 in the vicinity of Holliday junctions, although it is not yet known whether it has additional functions in these scenarios aside from migration of strands. Fujikane et al. (2006) found that there was an imperfect correlation between DNA structures that stimulated binding of Hjm, and structures that stimulated its ATP-ase activity. This could suggest that Hel308 possesses both ATP-dependent and non-ATP-dependent functions. It is possible that, in addition to its helicase activity, Hel308 may function as a target for other proteins or interact with them to bring about other effects.

Immobilised *Methanothermobacter thermautotrophicus* Hel308 has been shown to interact with RPA, both in the presence and absence of DNA (Woodman et al., 2011). This ability to associate with RPA was ablated by mutation of the middle arginine of the RxRAR motif in domain 5, although this may be due to disruption of the protein structure rather than the loss of a binding site. HELQ's helicase activity is likewise stimulated by RPA (Tafel et al., 2011). It has been suggested that RPA could play a role in recruiting Hel308 to ssDNA structures, although given the Hel308-DNA co-crystal structure discussed above, it is clearly not a requirement for helicase loading.

In-vivo immunoprecipitation assays in *P. furiosus* have shown that Hjm co-precipitates with PCNA (Fujikane et al., 2006). The final 20 residues of Hjm prove essential for this interaction. This sequence includes a putative PIP-box in the *Pyrococcus* homologues examined, but is absent in other families of archaea.

In addition to separating duplex strands, human HELQ is able to cause complementary ssDNA strands to anneal. This activity has also been observed in the F295A point mutant of *M. thermautotrophicus* Hel308 (Lever et al., 2023). The relevance and mechanism of this activity is not yet understood, but it is possible that the secondary DNA binding site in domain 5 posited by Richards et al. (2008) could bind a different ssDNA strand to that being translocated along, bringing the two into close proximity and thus encouraging their interaction.

1.5.2.2 1.4 Hel308 functions in *Haloferax volcanii*

In vivo, Hel308 functions as a "guardian of genome stability". Homologues in *Drosophila* were identified through mutations causing sensitivity to DNA damaging agents such as mitomycin C (McCaffrey et al., 2006). Mutations in the human homologue HELQ can be associated with increased risk of an extensive range of cancers, suggesting a potential role in tumour suppression; reviewed in Tang et al. (2023).

Hel308 is not essential in most archaea species, but has been found to be essential in *Sulfolobus tokodaii* (Hong et al., 2012) and *Sulfolobus islandicus* (Zheng et al., 2012). The *hel308* gene was originally thought to be essential in *H. volcanii*, as attempted deletion of this gene using standard techniques never yielded deletion pop-outs. However, further study later revealed that the stop codon of *hel308* overlaps with the start codon of another gene, *cgi*, which is a component of the KEOPS complex, and is itself essential (Naor et al., 2012). KEOPS (kinase, endopeptidase, and other proteins of small size) is conserved throughout eukarya and archaea, and its major function is thought to be modification of tRNAs, although it has also been implicated in homologous recombination and DSB repair (He et al., 2019). Previous *hel308* deletion attempts had inadvertently

deleted the start codon of this gene in the process of removing *hel308*. Design of *hel308* deletion vectors that leave the *cgi* gene intact have allowed production of *Δhel308 H. volcanii* cells, proving its inessentiality in this species (T. Allers, unpublished research).

H. volcanii cells lacking Hel308 show a marked growth defect, resulting in a doubling time of 5.25 hours compared to around 2.75 for wild-type cells (Lever, 2020). *Δhel308* strains also demonstrate increased recombination compared to wild type, and increased sensitivity to mitomycin C, suggesting a role for Hel308 in regulating homologous recombination responses to DNA damage.

A second Hel308 gene is present in *H. volcanii,* denoted *hel308b* (Lever, 2020). This gene appears to be the result of a gene duplication event, and resembles a truncated form of *hel308*, being 639 amino acids in length, as compared to the 827 amino acid full-length variant. As a result, domain 5 is entirely missing from *hel308b*. Deletion of this gene does not confer the susceptibility to DNA crosslinking seen in strains lacking *hel308*, but does affect recombination. This suggests a secondary role for *hel308b*, and, by extension, perhaps an additional role for *hel308*.

2. Materials and Methods

2.1 Materials

2.1.1 Plasmids

Table 2-1 Plasmids used in this study. Plasmids without notes were produced for this study.

Name	Use	Notes
p357	Empty vector bearing *pyrE2* marker	Produced by S. Haldenby
p409	Empty vector bearing *pyrE2* and *hdrB* markers	Produced by T. Allers
p1276	Deletion vector for *hel308* in *Haloferax volcanii*	Produced by T. Allers
p1335	Chromosomal replacement of *H. volcanii hel308* with *hel308-D145N*	Produced by T. Allers
p1337	*dam*- preparation of p1335	Produced by T. Allers
p1451	Source of the *p.tna*M3 promoter	Produced by H. Marriott
P1642	Chromosomal replacement of *H. volcanii hel308* with *hel308-F316A*	Produced by R. Gamble-Milner
p1647	*dam*- preparation of p1642	Produced by R. Gamble-Milner
p1854	Plasmid bearing *H. volcanii hel308* gene	Produced by B. Lever
p1988	Chromosomal replacement of *H. volcanii hel308* with *hel308-K53A*	Produced by B. Lever
p2015	*dam*- preparation of p1988	Produced by B. Lever
p2353	Vector used for *H. mediterranei* genome library	Produced by A. Dattani
p2482	Vector for *H. mediterranei mrr* deletion	Produced by A. Dattani
p2483	*dam*- preparation of p2482	Produced by A. Dattani
p2680	Plasmid bearing *H. mediterranei hel308* for complementation	Produced by A. Dattani
p2713	Empty vector for point mutant lethality experiment	Produced by MSci student O. Wood, under my supervision
p2714	Plasmid bearing *H. volcanii hel308* gene for screen verification	
p2715	Plasmid bearing *H. volcanii hel308* gene	Produced by A. Dattani
p2721	Plasmid bearing inducible *H. volcanii hel308*	Produced by MSci student O. Wood, under my supervision
p2736	Plasmid bearing inducible *H. volcanii hel308-D145N*	Produced by MSci student O. Wood, under my supervision
p2737	Plasmid bearing inducible *H. volcanii hel308-K53A*	Produced by MSci student O. Wood, under my supervision
p2757	Plasmid bearing *H. mediterranei hel308* gene for screen verification	
p2789	Plasmid bearing inducible *H. mediterranei hel308* gene	Produced by MSci student O. Wood, under my supervision
p2828	Plasmid bearing *H. mediterranei hel308-G68-T85del for complementation*	

p2861	Plasmid bearing *H. volcanii hel308* for complementation	
p2875	*H. mediterranei hel308* deletion vector	
p2876	*H. mediterranei hel308* and *cgi* deletion vector	
p2877	*trpA*-marked *H. mediterranei hel308* deletion vector	
p2878	*trpA*-marked *H. mediterranei hel308* and *cgi* deletion vector	
p2879	*H. mediterranei cgi* deletion vector	
p2880	*trpA*-marked *H. mediterranei cgi* deletion vector	
p2881	*dam*$^-$ preparation of p2877	
p2882	*dam*$^-$ preparation of p2880	
p2883	*dam*$^-$ preparation of p2878	

2.1.2 Strains

2.1.2.1 Haloferax volcanii strains

Table 2-2 Haloferax volcanii *strains used in this study. Strains without notes were produced for this study.*

Name	**Genotype**	**Notes**
H26	*ΔpyrE2*	Produced by T. Allers
H53	*ΔpyrE2* *ΔtrpA*	Produced by T. Allers
H77	*ΔtrpA*	Produced by T. Allers
H164	*ΔpyrE2* *ΔtrpA* *bgaHa-Bb* *leuB-Ag1*	Produced by T. Allers
H1208	*ΔpyrE2* *Δhdrb* *Δmrr*	Produced by T. Allers
H1391	*ΔpyrE2* *Δhel308*	Produced by T. Allers
H1393	*ΔpyrE2* *Δhel308::trpA*	Produced by T. Allers
H1555	*ΔpyrE2* *ΔtrpA* *hel308-D145N*	Produced by T. Allers
H1804	*ΔpyrE2* *ΔtrpA* *ΔoriC1* *ΔoriC2* *ΔoriC3* *Δori-pHV4*	Produced by K. Ptasinska
H2085	*ΔpyrE2* *ΔtrpA* *Δhel308*	Produced by D. Wallbank
H5074	*ΔpyrE2* *Δmrr* *ΔoriC1*	Produced by A. Dattani.

	ΔoriC2 *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3*	
H5107	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *ΔoriC1::[oripHV1 p.tnaA M3::orc10 LeuB+]*	Produced by A. Dattani.
H5366	*ΔpyrE2* *ΔtrpA* *ΔoriC1* *ΔoriC2* *ΔoriC3* *Δori-pHV4* *Δhel308::trpA*	Produced by A. Dattani

Parent: H53 (*ΔpyrE2, ΔtrpA*)

Name	**Genotype**	**Notes**
H5701	*Δhel308::trpA+::[hel308-k53A pyrE2+]*	

Parent: H164 (*ΔpyrE2, ΔtrpA, bgaHa-Bb, leuB-Ag1*)

Name	**Genotype**	**Notes**
H4361	*ΔpyrE2* *ΔtrpA* *bgaHa-Bb* *leuB-Ag1* *Δhel308*	Produced by B. Lever
H2400	*hel308-D145N*	Produced by R. Gamble-Milner.
H2397	*hel308-F316A*	Produced by R. Gamble-Milner.

Parent: H1208 (*ΔpyrE2, Δhdrb, Δmrr*)

Name	**Genotype**	**Notes**
H5452	*Hel308+*::[*Δhel308 pyrE2+]*	
H54553	*Δhel308*	

H5525	*[pyrE2+ hdrB+]*	Produced by MSci student O. Wood, under my supervision
H5524	*[p.tnaM3::hel308 pyrE2+ hdrB+]*	Produced by MSci student O. Wood, under my supervision
H5523	*[p.tnaM3::hel308-K53A pyrE2+ hdrB+]*	Produced by MSci student O. Wood, under my supervision
H5522	*[p.tnaM3::hel308-D145N pyrE2+ hdrB+]*	Produced by MSci student O. Wood, under my supervision

Parent: H1391 (*ΔpyrE2, Δhel308)*

Name	**Genotype**	**Notes**
H5533	*{pyrE2+}*	
H5608	*{p.tnaA::hel308+ pyrE2+ hdrb+}*	
H5609	*{p.tnaA::Hmedhel308+ pyrE2+ hdrb+}*	

Parent:H1393 (*ΔpyrE2, ΔtrpA, Δhel308::trpA+)*

Name	**Genotype**	**Notes**
H5634	*Δhel308:: trpA+::[hel308-K53A pyrE2+]*	
H5721	*hel308-K53A*	

Parent: H1804 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3, Δori-pHV4*)

Name	**Genotype**	**Notes**
H5366	*Δhel308:: trpA+*	Produced by A. Dattani

Parent: H5074 (*ΔpyrE2, Δmrr, ΔoriC1, ΔoriC2, ΔoriC3, ΔoripHV3, Δori-pHV4-2, ΔpHV1, ΔleuB::[pHV3], Δ3'-5'LeuB::pHV3*)

Name	**Genotype**	**Notes**
H5613	*{p.tnaA::hel308+ pyrE2+ hdrB+}*	
H5614	*{p.tnaA::Hmed hel308+ pyrE2+ hdrB+}*	

Parent: H5107 (*ΔpyrE2, Δmrr, ΔoriC1, ΔoriC2, ΔoriC3, ΔoripHV3, Δori-pHV4-2, ΔpHV1, ΔleuB::[pHV3], Δ3'-5'LeuB::pHV3, ΔoriC1::[oripHV1 p.tnaA M3::orc10 LeuB+]*)

Name	**Genotype**	**Notes**
H5526	*{hel308+ pyrE2+}*	
H5527	*{Hmedhel308+ pyrE2+}*	

Parent: H5366 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3, Δori-pHV4, Δhel308::trpA+*)

Name	**Genotype**	**Notes**
H5681	*Δhel308::trpA::[hel308-F316A pyrE2+]*	
H5683	*Δhel308::trpA::[hel308-D145N pyrE2+]*	

H5701	*Δhel308::trpA::[hel308-K53A pyrE2+]*	
H5734	*Δhel308::trpA::[Δhel308 pyrE2+]*	
H5714	*hel308-F316A*	
H5720	*hel308-D145N*	
H5725	*hel308-K53A*	
H5736	*Δhel308*	

Parent: H5452 (*ΔpyrE2, ΔhdrB, Δmrr, Hel308+*::[*Δhel308 pyrE2+]*)

Name	**Genotype**	**Notes**
H5453	*Δhel308*	

Parent: H5453 (*ΔpyrE2, ΔhdrB, Δmrr, Δhel308*)

Name	**Genotype**	**Notes**
H5542	*[pyrE2+ hdrB+]*	
H5543	*[p.tna*M3::*hel308 pyrE2+ hdrB+]*	
H5544	*[p.tna*M3::*hel308-D145N pyrE2+ hdrB+]*	
H5545	*[p.tna*M3::*hel308-K53A pyrE2+ hdrB+]*	

2.1.2.2 Haloferax mediterranei strains

Table 2-3 Haloferax mediterranei *strains used in this study. Strains without notes were produced for this study.*

Name	**Genotype**	**Notes**
H828	*ΔpyrE2* *ΔtrpA*	From M. Mevarech
H4676	*ΔpyrE2* *ΔtrpA* *ΔoriC1* *ΔoriC2* *ΔoriC3*	

Parent: H828 (*ΔpyrE2, ΔtrpA)*

Name	**Genotype**	**Notes**
H5222	*hel308+ ::[Δhel308::trpA+ pyrE2+]*	Produced by A. Dattani
H5682	*mrr::[Δmrr pyrE2]*	
H5739	*ΔpyrE2* *ΔtrpA* *hel308+::[Δhel308::trpA+ pyrE2+]*	
H5740	*ΔpyrE2* *ΔtrpA* *cgi+::[Δcgi::trpA+ pyrE2+]*	
H5741	*ΔpyrE2* *ΔtrpA* *hel308+::[Δhel308 Δcgi::trpA+ pyrE2+]*	

Parent: H4676 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3)*

Name	Genotype	Notes
H5742	*hel308+ ::[Δhel308::trpA+ pyrE2+]*	

Parent: H5222 (*ΔpyrE2, ΔtrpA, hel308::[Δhel308::trpA+ pyrE2+])*

Name	Genotype	Notes
H5344	*{mevR}*	Produced by A Dattani
H5580	*{Hmed hel308+ mevR}*	
H5594	*{Hmed hel308-G68-T85del+ mevR}*	
H5612	*{hel308+ mevR}*	

2.1.2.3 *E. coli* strains

Table 2-3 E. coli strains used in this study.

Name	Genotype	Notes
XL1-Blue MRF'	*endA1* *gyrA96 (NalR)* *lac [F'proAB lacIqZΔM15 tn10 (TetR)]* *Δ(mcrA)183* *Δ(mcrCB-hsdSMR-mrr)173* *recA1* *relA1* *supE44* *thi-1*	Standard cloning strain for blue/white screening using pBluescript based plasmids. Tetracycline resistant. Restriction endonuclease and recombination deficient. dam+. From Stratagene.
N2338 (GM121)	*F-* *ara-14* *dam-3* *dcm-6* *fhuA31* *galK2* *galT22* *hsdR3* *lacY1* *leu-6* *thi-1* *thr-1* *tsx-78*	*dam- dcm-* mutant for preparing DNA for *H. volcanii* transformations (Allers et al., 2004). From RG Lloyd.
E14	*F-* *thr-1* *ara-14* *leuB6* *Δ(gpt-proA)62* *lacY1* *tsx-33* *supE44* *galK2* *lamba-* *rac-* *hisG4 (Oc)*	*dam- strain.* From A Babic.

	rfbD1 *mgl-51* *rpsL31* *kdgK51* *xyl-5* *mtl-1* *argE3 (Oc)* *thi-1* *qsr'* *dam::Cm*	
E15	F- thr-1 ara-14 leuB6 Δ(gpt-proA)62 lacY1 tsx-33 supE44 galK2 lamba- rac- hisG4 (Oc) rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1 argE3 (Oc) thi-1 qsr' dam::Cm mutS::spect/strm	*dam- mutS- strain.* From A Babic.
E16	F- thr-1 ara-14 leuB6 Δ(gpt-proA)62 lacY1 tsx-33 supE44 galK2 lamba- rac- hisG4 (Oc) rfbD1 mgl-51 rpsL31 kdgK51 xyl-5 mtl-1	*dam- mutS- recA- strain.* From A Babic.

	argE3 (Oc) thi-1 qsr' dam::Cm mutS::spect/strm recA srl::Tn10	

2.1.3 1.3 Oligonucleotides

Oligonucleotides were synthesised by Eurofins MWG, Germany.

Table 2-4 Oligonucleotides used in this study. Mismatches with the template sequence are indicated by lower case letters.

Oligo code	**Full name**	**Sequence**	**Purpose**
o113	dlhr5F2	GCAAGGCGATTAAGTTGGGTAACG	Construction of p2828
o306	QR-XbaI	CCATCTAGAGGTCGATGGGTCGC	Colony PCR of *H. mediterranei hel308* deletion candidates
o867	hel308EcoR	AGGTAGTCGAGCACGCGGTCC	Construction of p2736, p2737, p2828
o1645	Hel308NdeI_F	GGGAGATcatATGCGAACTGCGGACCTGACGGGC	Construction of p2736 and p2737
o2264	HmedHel308DS_R_Kpn1	GCCGACGGTACCGAGCACGATTTTCGT	Colony PCR of *H. mediterranei hel308* deletion candidates
o2274	HMEStrepHel308_R	TCGCTTCGGaTcCGAGTACTCTCATTCGAAATCAC	Construction of p2789
o2370	hel308 ClaI Fw	CGCCATCGatCCGCCGGGCCGTAC	Construction of p2714 and p2858
o2371	hel308 ClaI Rv	AACGAATCGAtGTCCGAGACGGTC	Construction of p2714 and p2858
o2379	Hm_hel308_ApaI_Fw	CATTTTCCGCCGGGCCcTACATAGC	Construction of p2757
o2417	Hm_hel308_BamHI_Rv	AACGCGTCGgGaTCcGAGACGGTCG	Construction of p2757
o2464	Hm_hel308_NdeI_F2	GGGGATcatATGCGAACAGCGGACC	Construction of p2789
o2472	Hmed_hel308_16aa-del_F	cgtcgcacgcGGTGGGAAAGCACTCTACATCGTCCC	Construction of p2828

o2523	Hmedhel308-F1-XbaI	CTTGGCTCtagaTGGCACGATTACTACCC	Construction of *H. mediterranei Δhel308 and Δcgi vector*
o2524	Hmedhel308-F2-NdeI	GATTTCatATGAGAGTACTCGAAGCCG	Construction of *H. mediterranei Δhel308 vector*
o2525	Hmedhel308-F3-NdeI	GAGTCGcATatgCGGCTCGGCTGC	Construction of *H. mediterranei Δhel308 and Δcgi vector*
o2528	Hmedhel308-R3-NdeI	TGTTCGCATatgATCCCCCTTGGTCCG	Construction of *H. mediterranei Δhel308 and Δcgi vector*
o2529	Hmedhel308-R1-ApaI	TAGAACAGGGcCcAATAGTTGGTGG	Construction of *H. mediterranei hel308* and *cgi* deletion vectors
o2530	Hmedhel308-R2-NdeI-2	GGTCGCaTatGCTTCGAGTACTCTCATTCG	Construction of *H. mediterranei Δcgi vector*
o2535	Hmed_mrr_F	CGAAGACCAAGCGATGGCCC	Generation of probe for deletion of *H. mediterranei mrr* gene
o2536	Hmed_mrr_R	GCTTCTTCGTTGAGTCCCGCG	Generation of probe for deletion of *H. mediterranei mrr* gene

2.1.4 Chemicals and enzymes

All enzymes were purchased from New England Biolabs (NEB) and all chemicals from Sigma, unless otherwise stated. Enzymes were used following the manufacturer's instructions.

2.1.5 Media

2.1.5.1 Haloferax media and solutions

Media are sterilised by autoclaving at 121°C for 1 minute. Solid media are stored at 4°C in sealed bags to prevent dessication and crystallisation. Plates are dried for at least 30 minutes before use. Liquid media are stored in the dark at room temperature. Other solutions are stored at room temperature unless otherwise stated.

30% salt water (SW)
4 M NaCl,
148 mM $MgCl_2.6H_2O$,
122 mM $MgSO_4.7H_2O$,
94 mM KCl,
20mM Tris.HCl pH7.5.

18% salt water (SW)
Made by diluting 30% SW with distilled water.
3 mM $CaCl_2$ added after autoclaving.

Trace elements
1.82 mM $MnCl_2.4H_2O$,
1.53 mM $ZnSO_4.7H_2O$,
8.3 mM $FeSO_4.7H_2O$,
200 µM $CuSO_4.5H_2O$.
Filter sterilised and stored at 4°C.

Hv-min salts
0.4 M NH_4Cl,
0.25 M $CaCl_2$,
8% v/v of trace element solution.
Stored at 4°C.

Hv-min carbon source
10% DL-lactic acid Na_2 salt,
8% succinic acid Na_2
salt·$6H_2O$,
2% glycerol,
Adjust pH to 7.0 with NaOH.
Filter sterilised.

10x YPC
5% yeast extract (Difco),
1% peptone (Oxoid),
1% casamino acids,
17.6 mM KOH.
Not autoclaved; made immediately prior to use.

10x Cas

5% casamino acids,
17.6mM KOH.
Not autoclaved; made immediately prior to use.

Hv-Cas Salts
362 mM $CaCl_2$,
8.3% v/v trace elements,
615 µg/ml thiamine,
77 µg/ml biotin.

KPO_4 Buffer
308 mM K_2HPO_4,
192 mM KH_2PO_4,
Adjust pH to 7.0 with NaOH.

Hv-YPC agar
1.6% agar (Bacto),
18% SW
10% v/v 10xYPC
3mM $CaCl_2$
Microwaved without 10x YPC or $CaCl_2$ to dissolve agar. 10xYPC added, then autoclaved. $CaCl_2$ added prior to pouring, once cooled.

Hv-Cas agar
1.6% agar (Bacto),
18% SW
10% v/v 10xCas
0.84% v/v Hv-Cas salts
0.002% v/v of KPO_4 buffer (pH 7.0)
Microwaved without 10x Cas or KPO_4 buffer to dissolve agar. 10xCas added, then autoclaved. KPO_4 buffer added prior to pouring, once cooled.

Hv-Min broth
18% SW,
30 mM Tris·HCl pH 7.5,
2.5% Hv- Min carbon source,
1.2% Hv-Min Salts,
0.002% v/v of KPO_4 buffer (pH 7.0),
444 nM biotin,
2.5 µM thiamine.
18% SW and Tris.HCL pH 7.5 autoclaved first. Other components added once cool.

2.1.5.2 *Haloferax* media supplements

Supplement	**Abbreviation**	**Final concentration**
Uracil	Ura	50µg/ml
Thymidine	Thy	50µg/ml (+50µg/ml hypoxanthine when supplementing Hv-Cas and Hv-Min)
Tryptophan	Trp	50µg/ml

5-Fluoroorotic acid	5FOA	50µg/ml + (10µg/ml uracil)
Mevinolin	Mev	4µg/ml

2.1.5.3 *E. coli* media

Broth was sterilised by autoclaving and stored at room temperature. Solid media was sterilised by autoclaving and stored at 4°C. Plates were dried for at least 30 minutes prior to use.

LB (lysogeny broth)
1% tryptone (Bacto),
0.5% yeast extract (Difco),
170mM NaCl,
2nM NaOH,
Adjust pH to 7.0.

LB Agar
300ml LB, 1.5% agar

LB-amp
LB broth or agar plus ampicillin, to a final concentration of 50µg/ml. Ampicillin was added after autoclaving, once cool.

SOC broth
2% tryptone (Bacto),
0.5% yeast extract (Difco),
10 mM NaCl,
2.5 mM KCl,
10 mM $MgCl_2$,
10 mM $MgSO_4$,
20 mM glucose.

2.1.6 Other solutions

TE
10 mM Tris.
HCl pH 8.0,
1 mM EDTA

Sodium acetate
3 M NaAc pH 5.2, filter sterilised

2.2 Methods

2.2.1 General *E. coli* microbiology

Growth and Storage of *E. coli*

Cultures of *E. coli* grown on solid media were incubated overnight in a static incubator (LEEC) at 37°C. Small-scale liquid cultures (5 ml) were grown overnight in the same static incubator with 8 rpm rotation. Large-scale cultures (300 ml) were incubated overnight in an Innova 4330 floor-standing shaking incubator (New Brunswick Scientific) at 37°C with 150 rpm shaking.

For short-term storage all cultures were stored at 4°C. For long-term storage, 20% (v/v) glycerol was added to cultures (from 80% glycerol stock), mixed and flash frozen using dry ice. Frozen stocks were then stored at -80°C.

Preparation of *E. coli* competent cells

Five strains of *E. coli*, XL-Blue, N2338, E14, E15, and E16 were used to prepare electrocompetent *E. coli* cells.

A 5 ml culture was grown overnight in LB broth at 37°C with 8 rpm rotation. Cells were diluted 1/100 in LB broth. These were grown at 37°C to A650= 0.5-0.8. Cells were pelleted at 6000 x g for 12 minutes at 4°C. The supernatant was removed, and the pellet resuspended in an equal volume of ice-cold sterile 1 mM HEPES (pH 7.5). This process was repeated using two thirds volume 1 mM HEPES (pH 7.5), one third volume 1 mM HEPES (pH 7.5), 0.1 volume 1 mM HEPES (pH 7.5) + 10% glycerol and finally 0.001 volume 1 mM HEPES (pH 7.5) + 10% glycerol. Cells were aliquoted into 100µl aliquots, snap frozen on dry ice and stored at -80°C.

Transformation of *E. coli* by electroporation

1-2 µg of DNA in 4 µl of sterile dH2O was added to 40 µl of electrocompetent cells on ice. The DNA and cells were gently mixed and transferred to a pre- chilled sterile electroporation cuvette (1 mm electrode gap, GENEFLOW). The cuvette was placed in an *E. coli* gene pulser (BioRad) and pulsed at 1.8 kV. 1 ml of SOC was immediately added and samples were incubated at 37°C with 8 rpm rotation for 1 hour, to allow for recovery of the cells. 100µl of the culture was plated onto LB-Amp plates and incubated at 37°C overnight.

2.2.2 General *Haloferax* microbiology

Growth and storage of Haloferax

Cultures of Haloferax volcanii grown on solid media were incubated for 1-7 days in a static incubator (LEEC) at 45°C in a plastic bag to prevent drying. Small-scale liquid cultures (1-10 ml) were grown overnight in the same static incubator with 8 rpm rotation. Large-scale cultures (>50 ml) were incubated overnight in an Innova 4330 floor-standing shaking incubator (New Brunswick Scientific) at 45°C with 120 rpm shaking. For short-term storage, plates and cultures were stored at room temperature. For long-term storage, 20% (v/v) glycerol was added to cultures (from 80% glycerol 6% salt water stock), mixed and flash frozen using dry ice. Frozen stocks were then stored at -80°C.

Transformation of Haloferax via PEG600

Haloferax can be effectively transformed using PEG600 (Cline et al., 1989). Strains carrying the *mrr* gene (encoding the Mrr restriction endonuclease) degrade DNA bearing *dam* methylation patterns.

Plasmid DNA must therefore be passaged through an *E. coli* strain lacking the *dam* gene to remove methylation patterns. No such precautions need be taken with *Δmrr Haloferax* strains.

Buffers and solutions:
All solutions are filter sterilised using a 0.2 µm syringe filter (Sartorius). Unless stated otherwise, all centrifuge spins were at 3300 x g, 25°C in a swing-bucket rotor.

Buffered Spheroplasting Solution:
1 M NaCl,
27 mM KCl,
50 mM Tris.HCl pH 8.5,
15% sucrose.

Unbuffered Spheroplasting Solution:
1 M NaCl,
27 mM KCl,
15% sucrose,
pH 7.5.

Transforming DNA:
5 µl 0.5 M EDTA, pH 8.0,
15 µl unbuffered spheroplasting solution,
10 µl DNA (~1-2 µg) (or dH_2O for control)

60% PEG 600:
150 µl PEG 600,
100 µl unbuffered spheroplasting solution.

Spheroplast Dilution Solution:
23% SW,
15% sucrose,
37.5 mM $CaCl_2$.

YPC Regeneration Solution:
18% SW,
10% v/v 10xYPC,
15% sucrose,
30 mM CaCl2.

Cas Regeneration Solution:
18% SW,
10% v/v 10xCas,
15% sucrose,
30 mM CaCl2.

Transformation Dilution Solution:
18% SW,
15% sucrose,

30 mM CaCl2

5 ml of YPC was inoculated with 1-4 colonies and incubated for ~16 hours at 45°C with 8 rpm rotation. When the A650 = 0.6-0.8, cells were pelleted by centrifugation in a 15 ml round-bottomed tube. The supernatant was removed, and the pellet was gently resuspended in 2 ml buffered spheroplasting solution. Cells were transferred to 2 ml round-bottomed tube, pelleted again, and the supernatant was removed. Cells were resuspended in 400-800 µl buffered spheroplasting solution. 200 µl of this suspension was transferred to a fresh 2 ml tube per transformation. 20 µl of 0.5 M EDTA (pH 8) was added to the side of the tube, gently inverted and incubated at room temperature for 10 minutes, facilitating removal of the cells' S-layer. DNA for transformation was added in the same manner as EDTA and incubated for a further 5 minutes at room temperature. 250 µl of 60% polyethylene glycol 600 (PEG600) was added to the side of the tube and mixed by gentle rocking, before incubation at room temperature for 30 minutes. 1.5 ml of spheroplast dilution solution was added and mixed by gentle inversion. Following a two-minute incubation at room temperature, cells were pelleted by centrifugation. The pellet was then transferred whole to a sterile 4 ml tube containing 1 ml regeneration solution (YPC or Cas-based, as appropriate). To allow recovery, cells were incubated statically at 45°C for 90 minutes. Cells were then resuspended by tapping the tube and incubated at 45°C with 8 rpm rotation for 3-4 hours. Cells were transferred to a fresh 2 ml round-bottomed tube and pelleted by centrifugation. The cell pellet was gently resuspended in transformation dilution solution. Appropriate dilutions were made and 100 µl of chosen dilutions were plated on suitable selective media. Plates were incubated at 45°C until individual colonies can be detected.

2.2.3 DNA extraction

Plasmid extraction from *E. coli*

Plasmid DNA extraction from *E. coli* was performed using Macherey-Nagel NucleoSpin Plasmid (Mini) and NucleoBond Xtra (Midi) kits. Protocol was followed as described in the manufacturer's guidelines. For minipreps 2 ml *E. coli* cell culture (LB broth +Amp) was used and eluted in 30 µl elution buffer. For midipreps 300 ml *E. coli* cell culture (LB broth +Amp) was used.

Midipreps were eluted using isopropanol. The DNA was ethanol precipitated, resuspended in 200 µl of TE, and stored at -20°C.

Plasmid extraction from Haloferax

Buffers and solutions:

ST buffer
1 M NaCl,
20 mM Tris.HCl pH 7.5.

Macherey-Nagel Nucleospin (Mini) kits were used to isolate circular plasmid DNA from *Haloferax*. The low plasmid copy number and large amount of cellular debris in *Haloferax* requires the following amendments to the manufacturer's guidelines:

Starter culture: 30ml of cell culture was used for minipreps.

Resuspension step: Cell pellets were resuspended in the same total volume as *E. coli* minipreps. However, the cell pellets were initially resuspended in ½ volume of ST buffer supplemented with 50 mM of EDTA, then ½ volume of the standard resuspension buffer (Macherey-Nagel) was added to make up the total volume.

Chloroform extraction: Chloroform extraction steps were required to separate DNA from the cellular debris in order to avoid column blockage. An equal volume of isoamyl alcohol:chloroform mixture (1:24) was added to cell lysates and vortexed to mix. The samples were then centrifuged for 1 minute at 3300 x g. The aqueous top layer was transferred to a fresh tube and the chloroform extraction step was repeated until the cell debris at the interface was minimal. The aqueous top layer was then loaded onto the column.

Genomic DNA extraction from Haloferax by spooling

Buffers and solutions:

ST buffer
1 M NaCl,
20 mM Tris.HCl pH 7.5.

Lysis buffer
100 mM EDTA pH 8.0,
0.2% SDS.

A 5 ml culture of H. volcanii was grown overnight in Hv-YPC at 45ºC until A650 = 0.6-0.8. 2 ml of culture was transferred to a round-bottomed 2 ml tube and centrifuged at 3300 x g for 8 minutes at 25ºC. The supernatant was removed and the pellet resuspended in 200 µl of ST buffer. 200 µl of lysis solution was added, the tube mixed by inversion, and the cell lysate overlaid with 1 ml of 100% EtOH. DNA was spooled at the interface with a capillary tip until the DNA could be seen wrapped around the tip. The spool of DNA was washed twice in 100% EtOH, and excess EtOH allowed to drain from the DNA. The DNA was air-dried, resuspended in 500 µl of TE, and precipitated with 400 µl isopropanol and 50 µl sodium acetate. The samples were centrifuged at 11,000 x g for 5 minutes, washed in 1 ml 70% EtOH, and dried thoroughly. The pellet was resuspended in 100-500 µl of TE and stored at 4ºC.

2.2.4 Nucleic acid amplification

PCR amplification

Amplification of DNA was carried out using Q5 HotStart or OneTaq (NEB). Q5 was used for amplifications that required high fidelity. OneTaq was used for diagnostic amplifications. Reaction conditions are shown below. Reactions were carried out in a Techne TC-512 Thermocycler.

Table 2-5 Reaction components for PCR

Component	Q5 HotStart	OneTaq

dNTPs	200µM of each dNTP	200µM of each dNTP
Primers	0.5µM of each primer	0.5µM of each primer
Template DNA	1 ng – 1 µg genomic DNA 1 pg – 1 ng plasmid DNA	10ng template DNA
Buffer	1xQ5 reaction buffer 1xQ5 high GC enhancer	1xOneTaq GC buffer
Enzyme	0.02 U/µl Q5 HotStart DNA Polymerase	0.025 U/µl OneTaq DNA Polymerase

Table 2-6 Reaction conditions for PCR

Step	**Q5 HotStart**	**OneTaq**	
Initial denaturation	98°C, 30 seconds	94°C, 30 seconds	
Denaturation	98°C, 10 seconds	94°C, 30 seconds	30 cycles
Annealing	Tm°C, 20 seconds	Tm°C, 30 seconds	
Extension	72°C, 30 seconds/kb	68°C, 60 seconds/kb	
Final extension	72°C, 5 minutes	68°C, 5 minutes	

Annealing temperatures (Tm°C) for primers were calculated using the following equation (Howley et al., 1979).

$$81.5 + (16.6\log10[\mathrm{Na+}]) + (0.41\mathrm{x}\%\mathrm{GC}) - (100 - \%\mathrm{homology}) - (\frac{600}{length})$$

Touchdown PCR

Where primers do not share 100% homology to the template DNA (as is the case when introducing restriction sites or mutations), two annealing temperatures were calculated. Tm_S was based on the original homology percentage between primer and template, while Tm_E was calculated assuming 100% homology. The reaction temperature started at Tm_S and increased linearly to Tm_E across 10 cycles. The remaining 20 cycles used Tm_E as the annealing temperature.

Colony PCR

Colony PCRs were used for screening large numbers of colonies for a desired plasmid or chromosomal gene. *Haloferax* or *E. coli* colonies were grown on solid media and touched lightly with a sterile 200µl pipette tip, ensuring only a small number of cells were collected and the colony was not disturbed. The pipette tip was pipetted up and down in 100 µl of sterile dH2O to dislodge the cells. This was then boiled at 100ºC for 10 minutes to lyse the cells and immediately cooled on ice. 1 µl of the preparation was then used in a PCR with OneTaq DNA polymerase.

Overlap extension PCR

The introduction of point mutations within a gene requires complementary forward and reverse primers containing the desired mutation. PCRs to amplify either side of the desired mutation are conducted using one of these complementary primers and an appropriate external primer. The products from these two PCRs are then mixed in an approximate equal proportion to undergo a third PCR. During the first 10 cycles of the reaction no primers were present and the annealing

temperature is calculated based on the overlap of the two products. The external primers only were then added to the mix, which amplified the mutated region for a further 20 cycles.

Restriction digests

Restriction digests were carried out following manufacturer's instructions (NEB). For double digests NEB buffers were selected so that each enzyme had at least 75% activity. Plasmid DNA was digested for at least one hour. Genomic DNA and PCR products were digested for 16 hours.

Blunt-ending with Klenow

Should overhangs generated by restriction digests have been required to be blunt-ended, the ends were filled in using Klenow (NEB). Samples were incubated with 1 unit of Klenow per µg of DNA, 1 mM dNTPs and 1x NEB buffer 4 for 30 minutes at 25°C. The reaction was stopped by heat inactivation at 75°C for 20 minutes.

Dephosphorylation of vector DNA

To prevent self-ligation of vector DNA, Shrimp alkaline phosphatase (NEB) was used to remove 5' phosphate groups. Samples were incubated with 5 units of Shrimp alkaline phosphatase per µg of DNA and 1x Antarctic phosphatase buffer (or CutSmart Buffer, commonly used in restriction digests) for 30 minutes at 37°C. Phosphatase was heat inactivated at 65°C for 10 minutes.

Ligation of DNA

Ligations were performed using T4 DNA Ligase (NEB). For each µg of DNA, 5 units of ligase were used in a reaction with 1× T4 Ligase buffer. For vector:insert ligations, reactions contained a molar ratio of ~3:1 insert to vector DNA. Ligations were carried out at 15°C overnight, followed by ethanol precipitation and resuspension in 5 µl dH2O. This DNA could then be used for transformation.

Ethanol precipitation of DNA

To ethanol precipitate DNA, 2 volumes of 100% EtOH and 1/10 volume of 3 M sodium acetate (pH 5.2) were added to DNA and incubated at -20°C for at least 1 hour. Samples were centrifuged at 20,000 x g for 30 minutes at 4°C and the supernatant removed. Pellets were washed in 400 µl 70% EtOH followed by centrifugation at 20,000 x g for 10 minutes at 4°C. The supernatant was removed, and the pellets air-dried thoroughly before resuspension in sterile dH2O.

Nucleic acid purification

PCR products were purified using Macherey-Nagel DNA purification kits. Protocol was followed according to manufacturer's instructions. In these kits, DNA is bound pH-dependently to a silica membrane and is separated from contaminants (such as small oligonucleotides or proteins) by washing with ethanol. DNA was eluted in 30 µl of the provided elution buffer.

Nucleic acid Quantification

To determine the concentration and purity of DNA preparations the absorbance at 260nm and the 260:280nm absorbance ratio respectively were measured by a Nanodrop Microvolume Spectrophotometer (Thermo Fisher).

DNA sequencing

All DNA sequencing reactions and analysis were performed by the Biopolymer Synthesis and Analysis Unit, University of Nottingham. Sequencing was carried out using the dideoxy chain termination method (Sanger et al., 1977).

Agarose Gel Electrophoresis

Buffers and solutions:

TBE (Tris/Borate/EDTA)
89 mM Tris.HCl,
89 mM boric acid,
2mM EDTA.

TAE (Tris/Acetic Acid/EDTA)
40 mM Tris.HCl,
20mM acetic acid,
1 mM EDTA.

Gel loading dye (x5)
50 mM Tris·HCl,
100 mM EDTA,
15% Ficoll (w/v),
0.25% Bromophenol Blue (w/v),
0.25% Xylene Cyanol FF (w/v).

TBE buffer was used as standard practice for casting and running agarose gels. TAE buffer was used when high quality resolution and/or Southern blotting was required. Agarose gels were cast using agarose powder (SeaKem Lonza) and either TBE or TAE buffer. Gel loading dye was added to the DNA samples to give a final concentration of 1x. All samples and molecular markers, either 1 kb or 100 bp (both NEB), were loaded onto the gel. TBE gels (10 cm) were run at 110 V for ~1 hour. TAE gels (25 cm) were run overnight (16 hours) at 50 V with buffer circulation. For visualisation of bands, gels were stained with ethidium bromide at a final concentration of 0.5 µg/ml or SYBR Safe (Invitrogen) at a 0.5x final concentration. Gels used for DNA extraction were stained with SYBR Safe.

Agarose gel extraction and purification of DNA

To purify DNA from agarose gels while avoiding UV exposure, gels were only exposed to SYBR Safe stain. DNA was visualised using a Dark Reader (Clare Chemical Research). DNA was purified using the Macherey-Nagel DNA purification kit following manufacturer's guidelines.

2.2.5 Genetic manipulation of *Haloferax*

Generating a deletion construct

Deletion constructs are generated by the insertion of the upstream and downstream regions of the gene of interest into cloning vector p131 (or a derivative). p131 is a derivative of a standard *E. coli* cloning vector, pBluescript II SK+, whereby the H. volcanii uracil biosynthesis gene *pyrE2* (encoding orotate phosphoribosyl transferase) has been inserted (Allers et al., 2004).

Upstream and downstream regions were generated by PCR in two separate reactions: one to generate the upstream region (US) and one for the downstream region (DS). These PCRs were performed using genomic DNA as a template. External primers were designed to incorporate specified novel restriction sites within the product, specifically one of those compatible with the multiple cloning site (MCS) of vector p131 or derivative. Internal primers were designed with a BamHI site, giving a BamHI site at the site of gene deletion. This gives a product of the US and DS regions flanking a BamHI site, which allows ease of downstream cloning when adding a marker flanked by BamHI (e.g. *trpA* from p298). PCR products were cut with BamHI and the newly introduced external restriction site, as appropriate, and ligated into compatible sites within p131 or derivative. Plasmids were transformed into *E. coli* XL-1 Blue cells and plated onto LB+Amp. Selected colonies were grown and screened by diagnostic restriction digest to check for the presence and correct orientation of the insert. Once confirmed by digest, DNA was sequenced to check for the absence of any point mutations.

Generating a gene replacement construct

Gene replacement constructs were made by inserting the gene of interest between the upstream (US) and downstream (DS) regions of a deletion vector, replacing the marker gene where present. The exact protocols for this differ and are discussed in the relevant chapter.

Gene deletion and replacement in Haloferax

Gene deletion and replacement constructs were used to transform *ΔpyrE2* strains. Transformants are plated on Hv-Ca (+ additives as required) to select for the integration of the *pyrE2*-marked plasmid at the targeted gene locus (pop-in, *pyrE2*+). A pop-in colony was picked and restreaked onto the same selective media. Confirmation of the pop-in was performed by diagnostic PCR.

This pop-in was then used to set up the pop-out culture: a 5 ml Hv-YPC culture (or Hv-Min+ura, in the case of *Haloferax mediterranei*) was inoculated with this single pop-in colony and the culture was grown overnight until A650 ≈ 1. This culture was then diluted 1/500 into a fresh 5 ml culture and the growth and dilution were repeated again. This was repeated five times. When the culture reached A650 ≈ 1, the culture was diluted in 18% salt-water and was plated on Hv-Ca +5-FOA (+additives as required).

The relief in selection for uracil in the subsequent Hv-YPC overnights allows for the integrative plasmid and native gene to be lost by homologous recombination. This loss of pyrE2 can then be selected for using 5-FOA to select for pop-outs (pyrE2-). These potential pop-out candidates were restreaked onto selective media and screened for the desired genotype. Depending on the

orientation of the pop-out, resulting colonies will be either wild-type or deletion/replacement mutants. The pop-in/pop-out gene replacement method was developed by (Bitan-Banin et al., 2003) and a schematic can be seen in Figure 13, below.

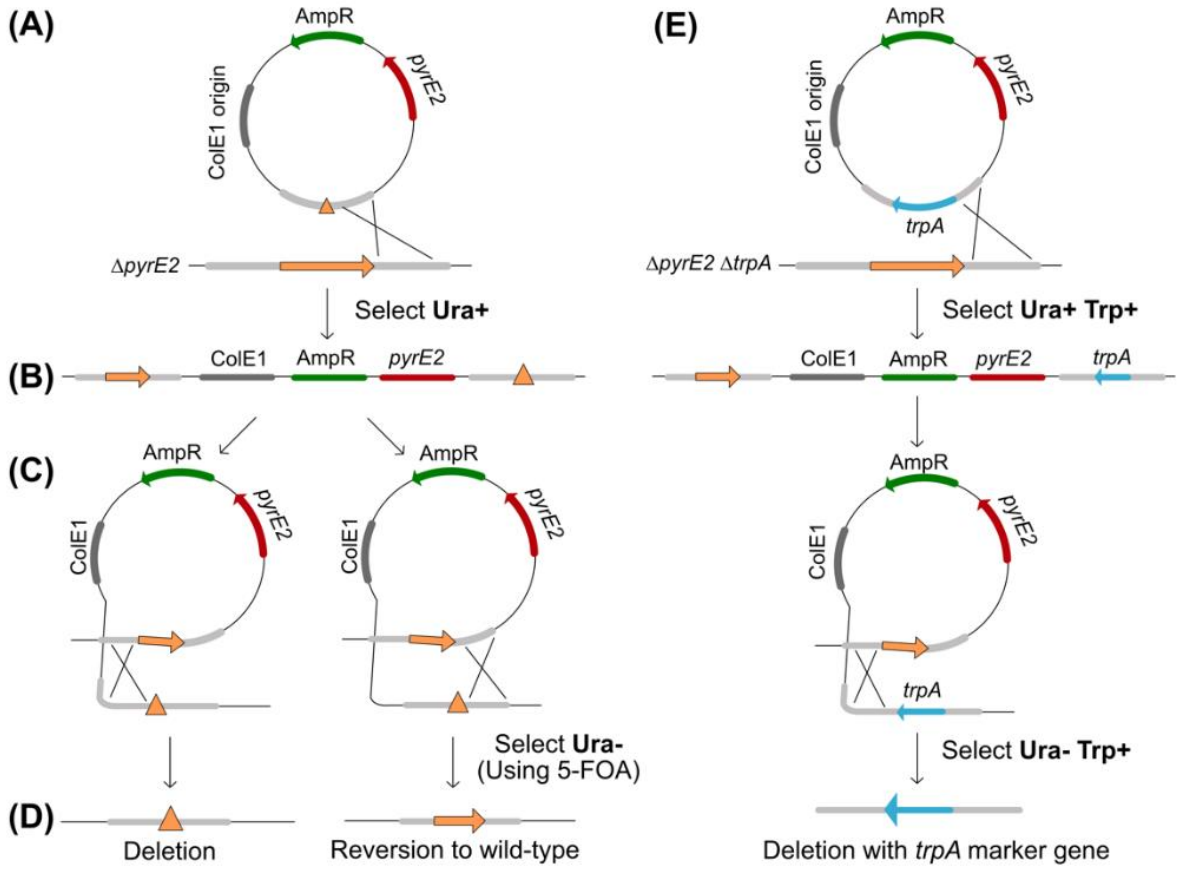


Figure 13. Gene deletion by the pop-in/pop-out method. (A) A ΔpyrE2 *strain is transformed with a* pyrE2*-marked deletion construct. (B) Pop-ins are selected by their ability to grow on media lacking uracil (ura+ phenotype). (C) Pop-out can be in the upstream (left) or downstream (right) orientation, resulting in the loss of the plasmid backbone (including pyrE2). The loss of* pyrE2 *in pop-outs is selected for by plating on 5-FOA. (D) The gene is deleted (left) or reverts to wild-type (right). (E) A trpA marker can be used in deletion constructs to allow for direct selection of deletion pop-out candidates. From Smith (2021).*

2.2.6 Genotype screening

Various methods exist for the screening of genotypes in *Haloferax*. Firstly, if the genotype being screened has a selectable phenotype (e.g., *ΔtrpA* strains cannot grow in the absence of tryptophan), then candidate colonies can be screened using selective media. However, the polyploid nature of *Haloferax* means genotypes can be merodiploid within a cell, where some chromosome copies may carry the desired alleles, while others retain wild-type alleles. Therefore, techniques need to ensure screening of all chromosome copies. For this reason, colony hybridisation and Southern blotting are used; these methods both require the denaturing of the total genomic DNA and transfer to a positively charged membrane by either colony lift or vacuum transfer.

Colony lift

Buffers and solutions:

20×SSPE:

3 M NaCl,

230 mM NaH_2PO_4,

32 mM EDTA, pH 7.4.

Denaturing Solution:
1.5 M NaCl,
0.5 M NaOH.

Neutralising Buffer:
1.5 M NaCl,
0.5 M Tris·HCl,
1 mM EDTA.

Colony lifts allow for screening of large numbers of colonies with relative ease. Candidate colonies were patched onto Hv-YPC using sterile wooden toothpicks and incubated at 45°C until growth (1-3 days). Patched colonies were lifted from the plate using circles of positively charged Nylon membrane (Amersham Hybond N+) following incubation on the plate for 1 minute. The membrane was transferred, colony side up, to Whatman paper soaked in 10% SDS for >5 minutes to lyse the cells. The membrane was then transferred to Whatman paper soaked in denaturing solution for >5 minutes to denature proteins and DNA. After this the membrane was transferred to Whatman paper soaked in neutralising solution for >5 minutes, which was then repeated. The membrane was then briefly washed for <30 seconds in 2 x SSPE before being thoroughly air-dried. DNA was crosslinked to the membrane with 120 mJ/cm2 UV.

Southern blot vacuum transfer

Southern blot utilises the same buffers and solutions listed for colony lift (above).

Purified *Haloferax* genomic DNA was digested with enzymes that cut either side of the region of interest. The resulting DNA fragments were separated on a 200 ml 0.75% TAE gel for 16 hours at 50 V, with buffer circulation. The gel was post-stained with 0.5 µg/ml ethidium bromide for 30 minutes and visualised. The gel-embedded DNA was acid nicked for 20 minutes in 0.25 M HCl, followed by washing for 10 minutes in dH2O. The DNA was then denatured in denaturing solution for 45 minutes. Membrane (BioRad Zeta-Probe GT or Amersham Hybond-XL) was soaked in dH2O for 5 minutes before equilibrating in denaturing solution for a further 2 minutes. Vacuum transfer was carried out using a Vacugene XL gel blotter and Vacugene Pump (Pharmacia Biotech) for 1 hour 15 minutes at 50 mBar. Following transfer, the membrane was washed briefly in 2 x SSPE for <30 seconds and air-dried before DNA was cross-linked with 120 mJ/cm2 UV.

Hybridisation

Buffers and solutions:

100 x Denhardt's Solution:
2% Ficoll 400,
2% PVP (Polyvinyl Pyrrolidone) 360,
2% BSA (Bovine Serum Albumin, Fraction V).

20 x SSPE:
3 M NaCl,

230 mM NaH2PO4,
32 mM EDTA, pH 7.4.

Pre-hybridisation solution:
6 x SSPE,
1% SDS.
5 x Denhardt's solution,
200 µg/ml salmon sperm DNA (Roche, boiled for 5 minutes at 100°C prior to addition).

Hybridisation Solution:
6 x SSPE,
1% SDS,
5% dextran sulfate (only added for Southern blots).

Wash 1
2x SSPE,
0.5% SDS

Wash 2
0.2x SSPE,
0.5% SDS

Membranes from either colony lifts or Southern blots were pre-hybridised for >3 hours at 65°C in 40 ml pre-hybridisation solution. Radiolabelled DNA probes were made with 50 ng of DNA and 0.74 MBq of [a-32P] dCTP (Perkin Elmer). DNA was denatured at 100ºC for 5 minutes then incubated with the radioisotope and HiPrime random priming mix (Roche) for 15-20 minutes at 37°C. The radiolabelled probe was then purified on a BioRad P-30 column and mixed with 10 mg/ml salmon sperm DNA, followed by denaturing at 100°C for 5 minutes and quenching on ice. For Southern blots 3 µl of 1 µg/ml 1 kb ladder or 3 µl of 1 µg/ml lambda DNA for Pulsed-Field Gel Southern blots was also included in the radiolabelling reaction. The pre-hybridisation solution was replaced with 30 ml of hybridisation solution, the probe DNA was added and then the membranes were incubated overnight at 65°C. The membranes were washed twice with 50ml wash 1 solution, once for 10 minutes and once for 30 minutes. This was followed by another two washes with wash 2 solution, both for 30 minutes. Membranes were air-dried before being wrapped in Saran wrap and exposed to a phosphorimager screen (Fujifilm BAS Cassette 2325) for >24 hours. The screen was scanned using a Molecular Dynamics STORM 840 scanner.

2.2.7 Phenotype screening

Tryptophan gradient plates

To generate a gradient of tryptophan across a plate, with the desired tryptophan concentration at one side fading to no tryptophan on the other side, plates were first poured with 17 ml Hv-Cas +Trp agar (of the desired concentration) on a 7° slant to form a tapered wedge (shown in Figure 14, below). Once set, the plate was placed flat and the wedge was covered with 43 ml Hv-Cas agar, lacking tryptophan. These were dried for at least 30 minutes prior to use.

5 ml cultures of H. volcanii strains were grown with 8 rpm rotation in Hv-Cas (+Trp where required) at 45°C until an A650 of 0.6-0.8. These were then diluted into fresh Hv-Cas and incubated at 45°C until an A650 of 1.0. Serial dilutions of the cultures in 18% SW to 10^{-5} were prepared. Autoclaved paintbrushes (The Range) were first wetted in 18% SW. Using a fresh paintbrush for each strain, the paintbrush was dipped into the diluted culture and then painted in one direction across the gradient plate. Using the same paintbrush dipped again in the diluted culture, a second line was painted over the first in the opposite direction. Once dry, the plates were incubated at 45°C for 5 days.

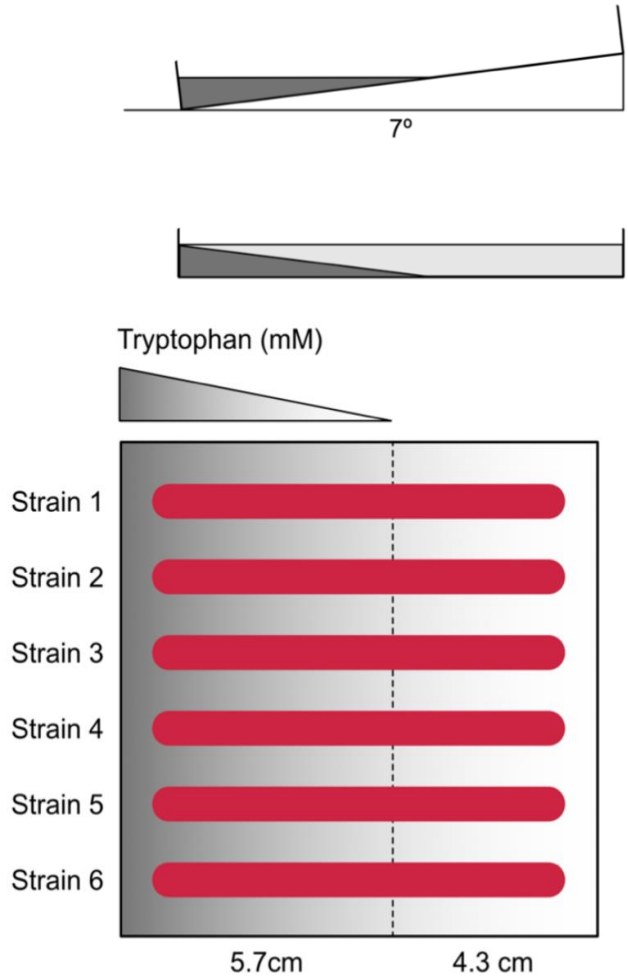


Figure 14. Protocol for tryptophan gradient plates. A wedge of Hv-Cas-trp agar (dark triangle) was poured into the square plate while on a 7° slant. Once set, the plate was placed flat and the wedge was covered with 43 ml Hv-Cas agar, lacking tryptophan. Strains are painted across the tryptophan gradient with autoclaved paintbrushes.

Standard growth assay

Standard growth rate in liquid media was determined using an Epoch 2 Microplate Spectrophotometer (BioTek). Cultures were prepared in 5 ml Hv-YPC or Hv-Cas broth and grown to mid-late exponential phase, which corresponds to an A650 of 0.4-0.8. These cultures were diluted and again grown to mid exponential phase. Serial dilutions of the cultures were made before loading 150 µl in duplicate, alongside appropriate blanks, to the wells of a 96 well microtiter plate (Corning). The outer wells along each edge of the 96-well plate were left blank, as these frequently crystalise, interfering with readings. Where drug treatment or other additive was required, this was added to

the media with which dilutions were made. For aphidicolin (VWR), a 1 mg/ml stock was diluted in DMSO and untreated cells received the same volume of DMSO as a control.

The plate was sealed around the edges with microporous tape (Boots UK Ltd) and incubated at 45°C with double orbital shaking at 1000 rpm for 72 hours in the spectrophotometer. Readings at A600 were taken every 15 minutes and converted to a 1 cm pathlength by dividing the raw A600 value by 0.14. The generation time was calculated by plotting the growth on a log2 scale and using the following equations:

$$G = \frac{t}{n}$$

$$n = \frac{\log b - log\mathrm{B}}{\log 2}$$

Where G=generation time, t=time, n=number of generations, b= end OD(A_{600}) B=start OD(A_{600}).

UV irradiation sensitivity assay

5 ml of Hv-YPC was inoculated with 1 colony and grown overnight at 45˚ C with 8 rpm rotation. The culture was diluted into 5 ml of fresh Hv-YPC and grown to an A650 of 0.4. A range of serial dilutions (10-1 – 10-8) of the cells in 18% saltwater were made and duplicate 20 µl samples were spotted onto Hv-YPC agar and allowed to air-dry. Plates were exposed to UV light at 254 nm, 1 J/m2/sec for varying amounts of time and shielded from visible light to prevent photo- reactivation DNA repair. Plates were incubated at 45°C for 4-7 days. During this period, colonies were counted, and survival fractions were calculated relative to a non-irradiated control.

3. Posited role of *hel308* in preventing origin-independent replication

3.1 Background

3.1.1 Origin-independent replication in *Haloferax*

3.1.1.1 Essentiality of origins in *Haloferax volcanii*

Not all DNA replication is initiated by origins. In *Haloferax volcanii*, not only is each origin individually non-essential, but cells remain viable even in the absence of all origins (Hawkins et al., 2013). While the deletion of any single origin from the chromosome typically confers some growth defect, cells lacking the three most active origins grew 5.5% faster than wild type, and cells in which all four chromosomal origins had been deleted were found to grow 7.5% faster than wild-type cells. This would appear to contradict previous reports (Ludt & Soppa, 2018) that some *Haloferax orc* genes are essential to cell function. However, they may still play other roles within the cell; for example, Orc proteins have been implicated in regulating DNA damage response in *Sulfolobus* species (Sun et al., 2018).

In cells without chromosomal origins, replication profiles generated by deep sequencing showed no well-defined peaks, suggesting that initiation of replication was occurring at diffused sites across the genome (Hawkins et al., 2013). While ORC proteins have been observed binding to other dispersed, non-specific sites around the genome in human cells (Vashee et al., 2003), this is thought not to be the case in *Haloferax.* All *orc* genes can be deleted from *H. volcanii*, producing the same replication profile seen here (Ausiannikava, unpublished data).

As the recombinase protein RadA becomes essential in origin-less cells, it is inferred that origin-independent replication is dependent on recombination, possibly using D-loops or R-loops to prime replication machinery (Hawkins et al., 2013).

3.1.1.2 Mechanisms of origin-independent replication

Origin-independent replication, including transcription-induced replication (TIR) and recombination-dependent replication, including break-induced replication (BIR) is well known in both bacteria and yeast, and has also been inferred in mammalian cells (Anand et al., 2013). However, BIR in yeast has been shown to be much more mutagenic than canonical DNA replication (Deem et al., 2011), which could explain the persistence of origins in the archaea despite the growth advantage conferred by their absence. Rate of mutation in origin-deleted *Haloferax* has not been explored in depth, but the rate of SNP generation in origin-less strains is not significantly different to wild-type (Hawkins et al., 2013).

BIR is a variety of recombination-dependent replication that arises from repair of a double strand break. Usually, during HR repair, the two broken DNA ends would both invade a homologous duplex, synthesise new DNA, and ligate to form a repaired sequence. However, in BIR, only one end of the double strand break successfully invades the homologous strand. Synthesis beginning at this invading strand functions much like a leading strand, forming a migrating D-loop that leaves the other duplex strand displaced, which can increase risk of DNA damage and mutation (Kramara et al., 2018). The newly synthesised strand can then be used as a template for its counterpart strand These are summarised in Figure 15, below.

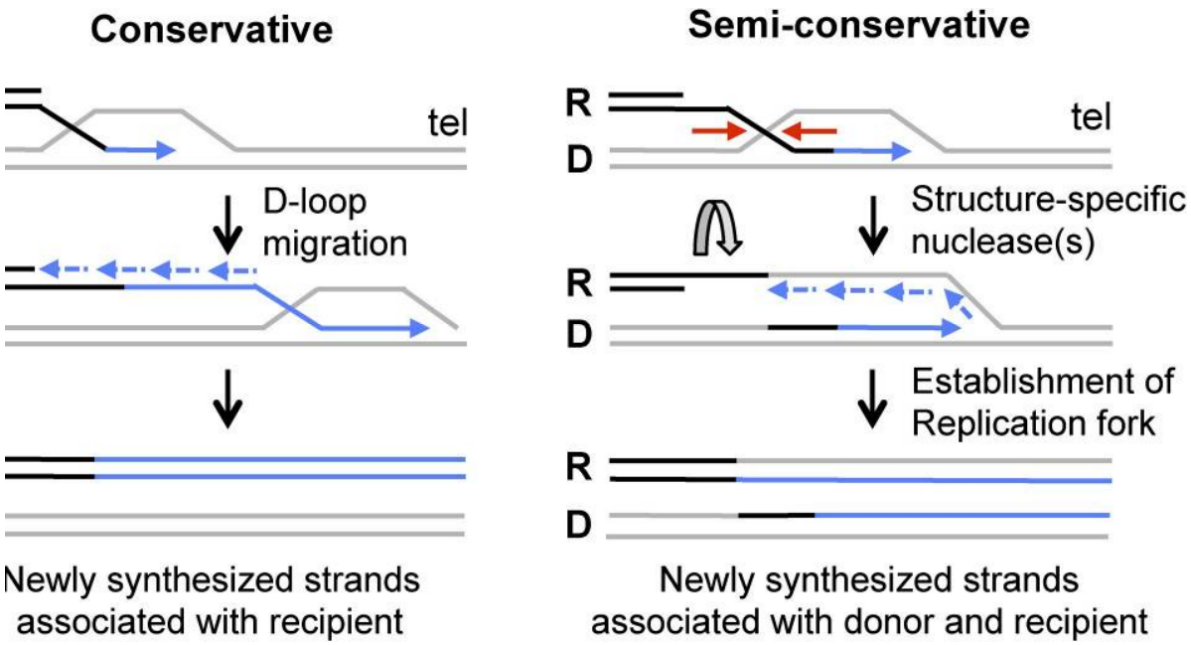


Figure 15. Generalised diagram of break-induced replication showing conservative and semi-conservative inheritance. Of these two, conservative inheritance is thought to be more common. R=recipient chromosome, D=donor chromosome. From Donnianni & Symington (2013).

The majority of study into BIR has been in bacteria, viruses, and eukaryotes such as yeast (Kramara et al., 2018), so the exact mechanisms at play in the archaea are subject to some conjecture. However, high ploidy organisms have the potential to be much better suited to this mode of replication. Given the greater number of homologous sequences available, the chances of the two ends of a double-strand break finding different partner molecules is much higher than it would be in diploid organisms. This means that each double-strand break repaired with HR could potentially prime two different replication events for that replicon. It is also possible for these DNA synthesis events to produce more double-strand breaks, as nicks encountered by replication forks are a common source of this kind of DNA damage (Malkova & Ira, 2013).

Treatment of origin-deleted *Haloferax* cells with the antimicrobial agent aphidicolin revealed another difference between the behaviour of these cells and wild-type. It was observed that these cells demonstrated resistance to aphidicolin, and that the more origins had been deleted, the stronger that resistance became (Smith, 2021). A growth assay showing the impact of aphidicolin is shown in Figure 16, below.

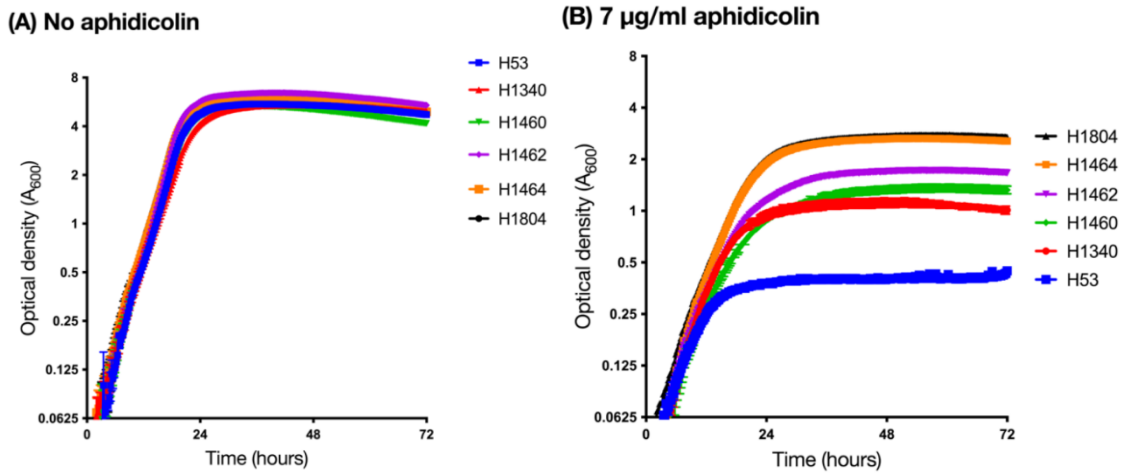


Figure 16. Impact of aphidicolin on Haloferax *strains. The greater the number of origins deleted, the greater the resistance to aphidicolin treatment. Genotypes of strains used are shown in the table. From Smith (2021).*

Strain	Genotype
H53	*ΔpyrE2, ΔtrpA,*
H1340	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2*
H1460	*ΔpyrE2, ΔtrpA, ΔoriC2, ΔoriC3*
H1463	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC3*
H1464	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*
H1804	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3, Δori-pHV4-2*

Aphidicolin is a DNA polymerase inhibitor that is specific to family B DNA polymerases – it therefore has no effect on the archaea-specific family D polymerase. It can be inferred from this result that origin-deleted cells make use of this polymerase more frequently than their wild-type peers, which could suggest preferential usage of PolD during DNA repair and recombination pathways, and a more central role of PolB in canonical origin-dependent replication.

In the oligoploid euryarchaeon *Thermococcus kodakarensis*, origins are non-essential (Gehring et al., 2017). However, replication profiles from this species indicate that origins are not used, even in wild-type strains. Furthermore, in this species PolD is essential, and PolB inessential, suggesting that PolD is the dominant DNA polymerase under normal growth conditions (Cubonová et al., 2013). *ΔpolB* strains suffer no growth defect, but are more susceptible to UV damage. This lack of growth defect demonstrates that PolD is capable of both leading and lagging strand synthesis, and can replicate the genome without the need for origins. The recombination factors *radA* and *radB* are also essential in wild-type *T. kodakarensis* (Gehring et al., 2017), suggesting that recombination-dependent replication may be the default mode of genome replication in this species. These findings lend further weight to the theory that PolB is the dominant DNA polymerase used in origin-dependent replication, while PolD may be the dominant DNA polymerase used in recombination-dependent replication.

As a final aside to the topic of origin essentiality in *H. volcanii*, it should be noted that the origin of the mini-chromosome pHV3 cannot be deleted, while those of pHV1 and pHV4 can (Marriott, 2018). Interestingly, deletion of the pHV1 origin does not result in the loss of pHV1 (Norais et al., 2007). As pHV3 contains several essential genes, it can be assumed that the lethality in this situation is caused by failure to replicate the plasmid, resulting in a lack of these essential genes. While it is not

immediately clear why replication of pHV1 is possible without origins, but not pHV3, it has been suggested that this may be due to the comparative transcriptional activities of these two structures. pHV3 is subject to generally low levels of transcription (Laass et al., 2019), so mechanisms of transcription-induced replication may have insufficient opportunities to initiate replication of this plasmid.

3.1.2 Origin independent replication in *H. mediterranei*

Haloferax mediterranei also bears three origins on its circular chromosome. Deletion of all three chromosomal origins is possible, but results in activation of an otherwise dormant fourth origin. Studies have shown that, while each of these four origins are individually non-essential, and deletion of any three is tolerated (with a corresponding decrease in growth rate), cells with origin-less main chromosomes cannot be produced (Yang et al., 2015).

This may be due to the lower recombination rate of *H. mediterranei*, which has recently been shown to be around 5 times lower than that of *H. volcanii* (Dattani, Sharon, et al., 2022). As *H. volcanii* cells without chromosomal origins have been shown to be dependent on recombination for DNA replication, it is possible that the lower recombination rates observed in *H. mediterranei* are not sufficient to meet the demands for genome replication in these circumstances.

As these two species are closely related and highly similar (many *H. volcanii* genetic tools function well in *H. mediterranei*), it is unexpected that they would behave dissimilarly in this regard. Given this point of contrast between two otherwise similar organisms, this potentially offers a window to explore what makes OIR possible or impossible. It can be posited that either *H. volcanii* possesses some factor that makes OIR possible, which *H. mediterranei* lacks, or *H. mediterranei* possesses some factor that prevents OIR, which *H. volcanii* lacks.

3.1.3 Prediction of origin essentiality

A bioinformatic tool was developed to predict whether species require origins or not, based on the qualities of their genomes (McCulloch, 2021). A range of features were considered, including GC skew profiles, Z-curves, co-orientation of core genes, and linkage of origins to replication initiator genes. These will be discussed briefly for context.

GC skew is a product of differing mutational pressures on leading and lagging strands, resulting in a bias of nucleotides present on a given strand. Leading strands contain more G and T, while lagging strands contain more C and A. This is quantified using sliding windows, typically using the formula $\frac{(G-C)}{(G+C)}$, and therefore ranges from 1 (in which C=0) to -1 (in which G=0). Near active origins, adjacent strands of the genome are reliably replicated as either leading or lagging strands; the effect diminishes with distance from the origin. At the origin (and terminator, if present), a sudden transition is seen in GC skew, which has previously been used to predict the location of origins in different organisms (Frank & Lobry, 2000). In the context of the bioinformatic tool, GC skew is expected to be stronger in species which are consistently using their origins, and noisier in species which make use of origin-independent replication mechanisms, as these may not initiate replication at specific sites of the genome. This “noise”, as a signal-to-noise ratio, can be calculated through fast

Fourier transformation of the data (Arakawa & Tomita, 2007). Example GC skew profiles generated in the course of the predictive tool development are shown in figure 17, below.

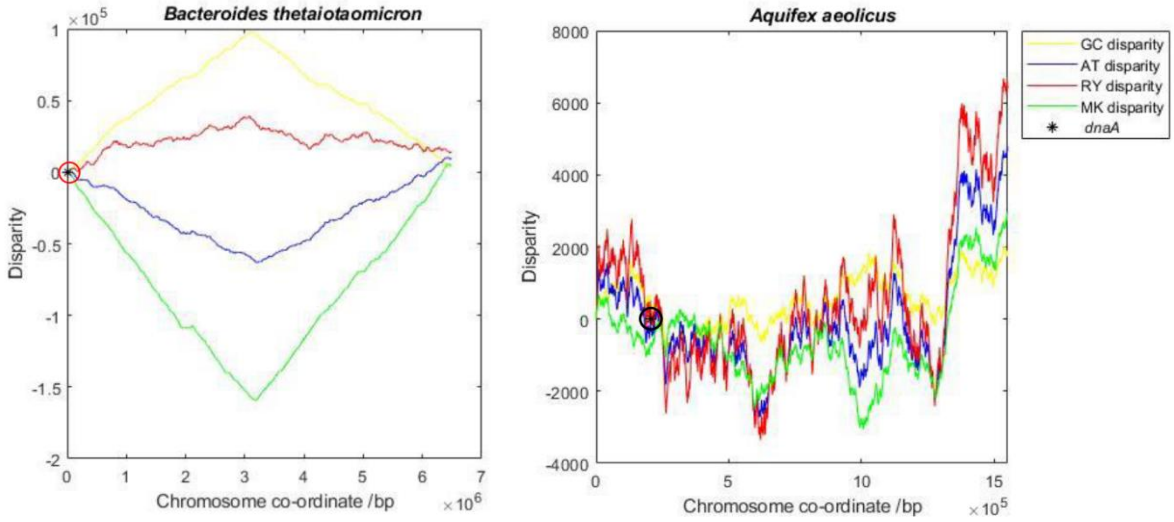


*Figure 17. Example GC skew profiles, showing one "clean" profile and one "noisy" profile. Nucleotide disparity types are indicated by the key; position of origin-associated genes are indicated by the *. Taken from McCulloch (2021)*

The concept of GC skew has been further developed into the Z-curve. In Z-curves, the cumulative frequency of each base type is represented as one dimension of a three-dimensional graph. Typically, the three axes represent the relative abundance of purines versus pyrimidines (A and G versus T and C), keto versus amino (A and C versus T and G), and weak versus strong hydrogen bonding (A and T versus C and G) (Zhang & Zhang, 2005). As this is cumulative frequency and does not utilise sliding windows, it allows finer resolution of skew in each dimension. As a side effect, the full genome sequence can be reconstructed from examination of the Z-curve. As with GC skew, points of sudden transition in can be used to infer the position of origins; noisier profiles may be correlated with reduced reliance on origins.

Positioning of genes annotated as involved in information storage and processing were also factored into the analysis. These tend to be clustered around origins in origin-dependent species. In addition, co-orientation of core genes (based on arCOG classification) is also a significant indicator of origin usage, as it helps avoid collision of transcription complexes with replication forks. Presence of DNA polymerases close to origins was also a weighted factor used in the screen (McCulloch, 2021).

The various factors were weighted in accordance with principal component analysis, and the finished bioinformatic tool applied to 86 bacterial and archaeal genomes. Some of these were known to be dependent on origins, some were known to be capable of origin-independent replication, and many were unknown in terms of origin essentiality. The output of the tool is a two-dimensional graph in which genomes predicted to be origin-dependent and origin-independent are physically separated, as shown in Figure 18, below.

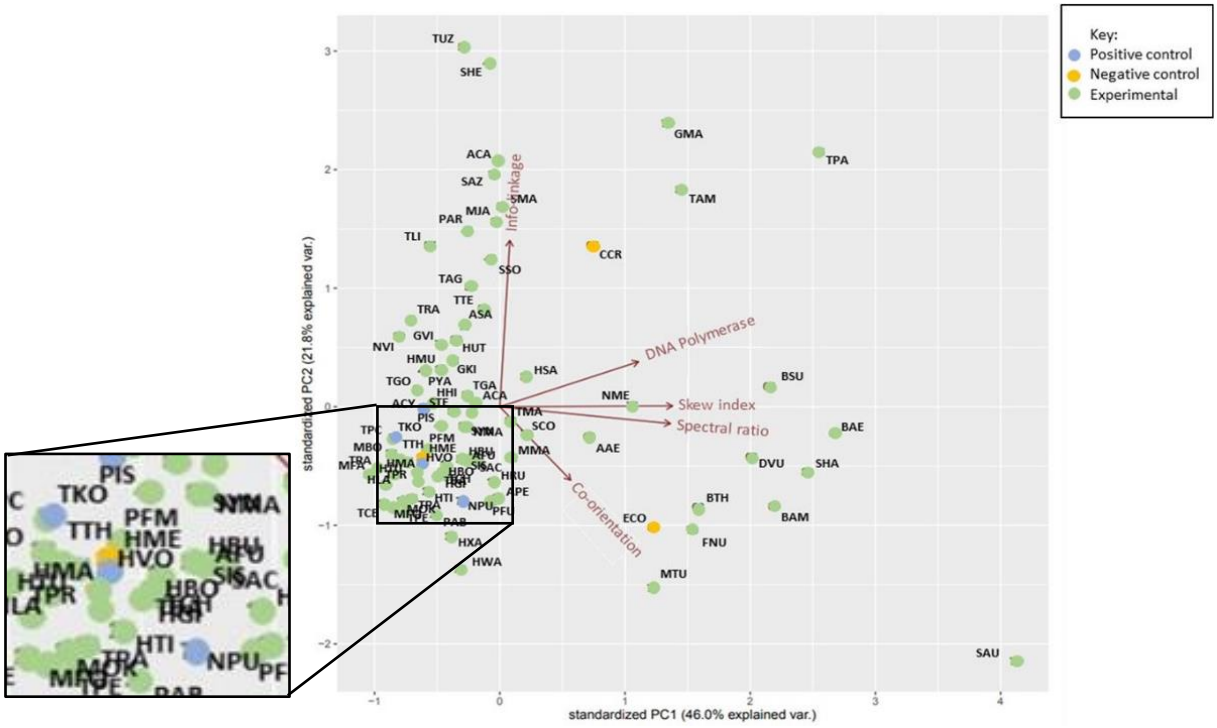


Figure 18. Output of the predictive bioinformatic tool. Despite its inclusion as a negative control (for origin-independent replication), H. mediterranei *(here denoted HME) is consistently placed adjacent to* H. volcanii *(here denoted HVO), in a region close to other species capable of origin-independent replication (blue dots). From McCulloch (2021).*

Despite some variations in weighting of the various factors, *Haloferax volcanii* and *mediterranei* were reliably placed in close proximity by the tool, in a neighbourhood containing other origin-independent species. While the proximity of the two *Haloferax* species is altogether not unexpected, given the high degree of similarity between them, it is notable that this results in *H. mediterranei* being consistently placed close to other species which are capable of origin-independent replication. This can perhaps be taken as an indication that it is *H. mediterranei* that possesses an inhibitor of origin-independent replication, rather than *H. volcanii* possessing an enabling factor of origin-independent replication. However, it is equally possible that this could be the result of *H. mediterranei* having recently lost some component that makes origin-independent replication possible, instead.

3.1.4 Screen for *H. mediterranei* anti-OIR factors

Based on the bioinformatic analysis, a screen was undertaken in the lab which aimed to identify factors in *H. mediterranei* which were capable of preventing origin-independent replication in *H. volcanii* (Dattani, unpublished work). A specialised *H. volcanii* strain was prepared for this screen. For ease of reading, the steps involved in this strain construction are not discussed in the order they were performed.

The three chromosomal *H. volcanii* origins were deleted (oriC1,2 and 3), as well as the origin from the pHV4 plasmid, which is incorporated into the main chromosome in the lab strain (Hawkins et al., 2013). The remaining *H. volcanii* plasmids, pHV1 and pHV3, also needed to be considered, since these bear their own origins which could potentially interfere with the screen. pHV1 was cured from the strain, as this plasmid carries no essential genes; its loss can therefore be tolerated by the cell (Marriott, 2018). pVH3, however, does contain some essential genes, and deletion of its origin has been found to be impossible. This plasmid was therefore forced to recombine onto the main chromosome, where it would be replicated along with the rest of the chromosome.

Recombination of pHV3 onto the chromosome was achieved through production of two strains bearing mutant *leuB* alleles in otherwise identical backgrounds. In one strain, a *leuB* gene bearing a 5' truncation was incorporated into pHV3 through pop-in, replacing the *adh2* gene. Deletion of *adh2* has been previously shown to cause no ill effects (Timpson et al., 2013). In the other strain, a *leuB* gene bearing a 3' truncation was incorporated onto the chromosome. Neither of these *leuB* genes produce functional enzymes. These two strains were then mated together, resulting in hybrid cells bearing both genomes. Recombination events between the two *leuB* alleles results in pHV3 being "popped in" to the chromosome, resulting in one full-length *leuB* allele and one bearing both truncations. Such strains can be identified by their restored leucine biosynthesis (Marriott, 2018). This is shown in Figure 19, below.

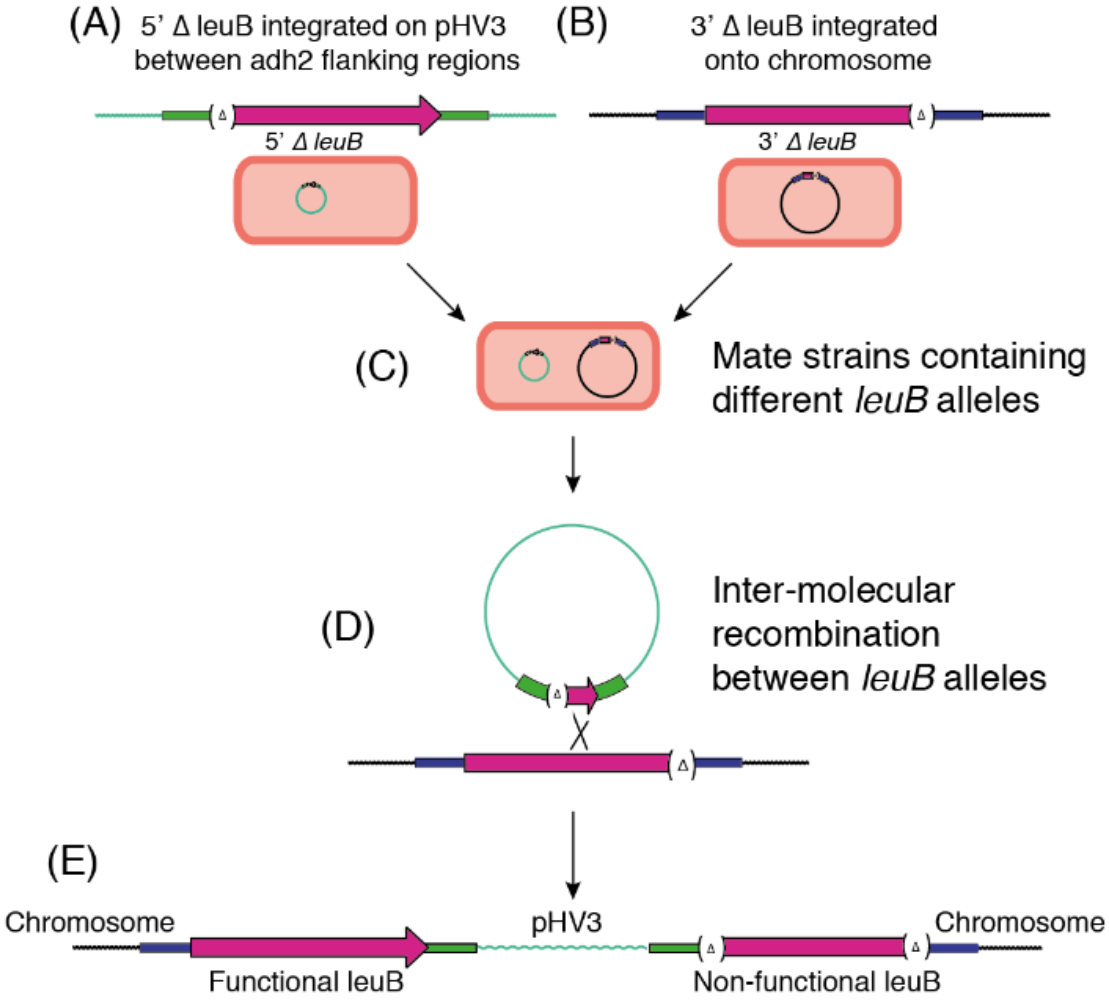


Figure 19. Incorporation of pHV3 into the main chromosome. (A) A strain containing the mutant leuBΔ5' *allele on pHV3 at the* adh2 *locus (B) and a strain containing the mutant* leuBΔ3' *allele on the main chromosome (C) are mated. (D) Intermolecular recombination occurs between the pHV3 mega-plasmid and the main chromosome (E) to produce a strain with pHV3 integrated on the chromosome flanked by a functional* leuB *and non-functional* leuBΔ5'Δ3' *allele. Figure from Marriott (2018)*

Once this event was confirmed (by Southern blot), the additional *leuB* gene was deleted through a linear transformation, to prevent further recombination "popping out" the integrated chromosome. As pHV3 would now be replicated as part of the main chromosome, its origin could also be deleted, resulting in a *H. volcanii* strain completely without origins. The origin on the small plasmid pHV2

need not be considered, as this plasmid has been cured from the standard lab strain (Norais, Hawkins et al. 2007).

Finally, at the position of the now-deleted oriC1, oripHV1 was inserted into the chromosome, alongside a copy of *orc10* under the control of the *p.tnaA* promoter. This origin was selected as it is thought to be a dominant origin within the cell, being the only candidate identified during a screen for autonomously-replicating sequences, and appearing multiple times (Norais et al., 2007). In addition, this origin appears to be under the control of a singular Orc protein. Unlike other *H. volcanii* Orc proteins, Orc10's activity is specific to the pHV1 origin (Ausiannikava, unpublished data), and this gene being placed under control of the tryptophan-inducible promoter allows the strain to be manually switched between two modes of DNA replication. In the presence of tryptophan, Orc10 will be expressed and will activate the pHV1 origin, replicating the chromosome through origin-dependent replication. In the absence of tryptophan, the origin will not be activated, and the genome will be replicated by origin-independent replication. This is summarised in Figure 20, below. This strain is referred to as H5107.

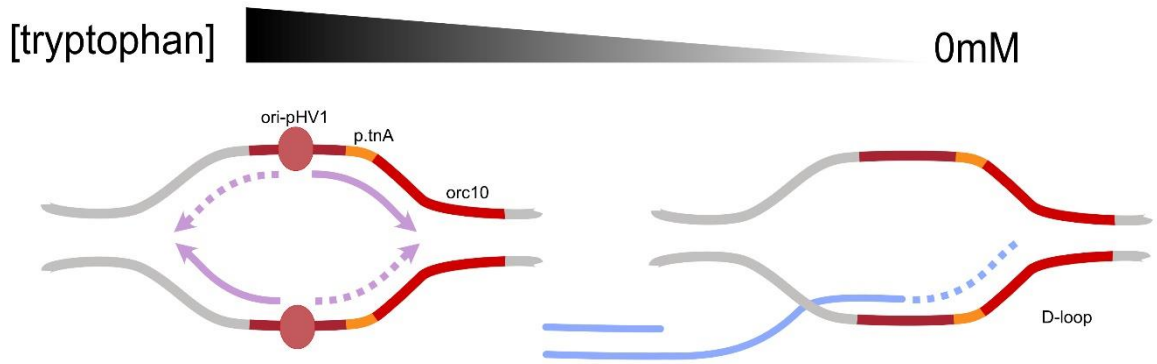


Figure 20. Schematic demonstrating origin usage under different tryptophan conditions. In the presence of tryptophan, expression of orc10 coordinates origin-dependent replication of the chromosome. In the absence of tryptophan, a lack of origin activation results in usage of recombination-dependent replication mechanisms. (Dattani, unpublished work)

In order to determine whether any components of the *H. mediterranei* genome could prevent origin-independent replication in *H. volcanii*, a *H. mediterranei* genome library was produced. Whole-genome DNA samples were taken from H826 through DNA spooling. This strain is *ΔpyrE2*, but otherwise wild-type *H. mediterranei*. The whole-genome samples were then subject to a partial digest with AciI. This restriction enzyme cuts at C^CGG, a very common sequence in halophile genomes, which have a high GC content in order to stabilise DNA in their high-salt intracellular environment (Paul et al., 2008). Following the partial digest, the results were visualised on an agarose gel and the 3kb-5kb fragments were excised, purified, and inserted into a vector bearing the oriC1 origin and *pyrE2* marker (p2353). The genome library was then transformed into H5107, and transformants plated on casamino acid-based plates with added tryptophan. Only cells which had been successfully transformed would be able to grow on this media, which lacks uracil. The presence of uracil in the media also ensures that all colonies are undergoing origin-dependent genome replication.

Around 5000 colonies from these plates were then replica plated onto plates with and without tryptophan. If the added plasmid bore a gene which interfered with origin-independent replication, the colony would not grow in the absence of tryptophan.

Of the 5000 colonies replica-plated, 6 exhibited tryptophan-dependent growth. Some of these candidates were assumed to be false positives; one contained an almost-complete *gloB* gene, a type II glyoxalase involved in glutathione metabolism, while another contained a uracil permease gene, *uraA*. As these genes are not thought to be pertinent to genome replication, they were dismissed as artefacts of the screening mechanism; particularly *uraA*, as uracil biosynthesis via *pyrE2* was used as a selective marker in the screen.

The more interesting candidate was one that contained two complete genes - ferredoxin and *hel308* helicase. As *hel308* is known to be involved in the mechanisms of recombination and DNA repair, this was considered the most promising lead of the available candidates.

Comparison of the *hel308* genes of *Haloferax volcanii* and *Haloferax mediterranei* shows that the genes are moderately well conserved; in fact, there is 88% amino acid sequence homology between the two. However, there is an 18 amino acid insert present in the *H. mediterranei* Hel308 sequence in domain 1 (shown in figure 21, below). Protein structure predictors seem to routinely place this 18-aa insert region on the exterior of the protein, so it is possible that it may be a site of a yet unidentified protein-protein interaction. This sequence is proline rich and is absent from most other *hel308* homologues, although it is present in *Haloferax larsenii* (Dattani et al, unpublished work). As *H. larsenii* has been studied less extensively than other members of this genus, it is not known whether its origins are essential or not.

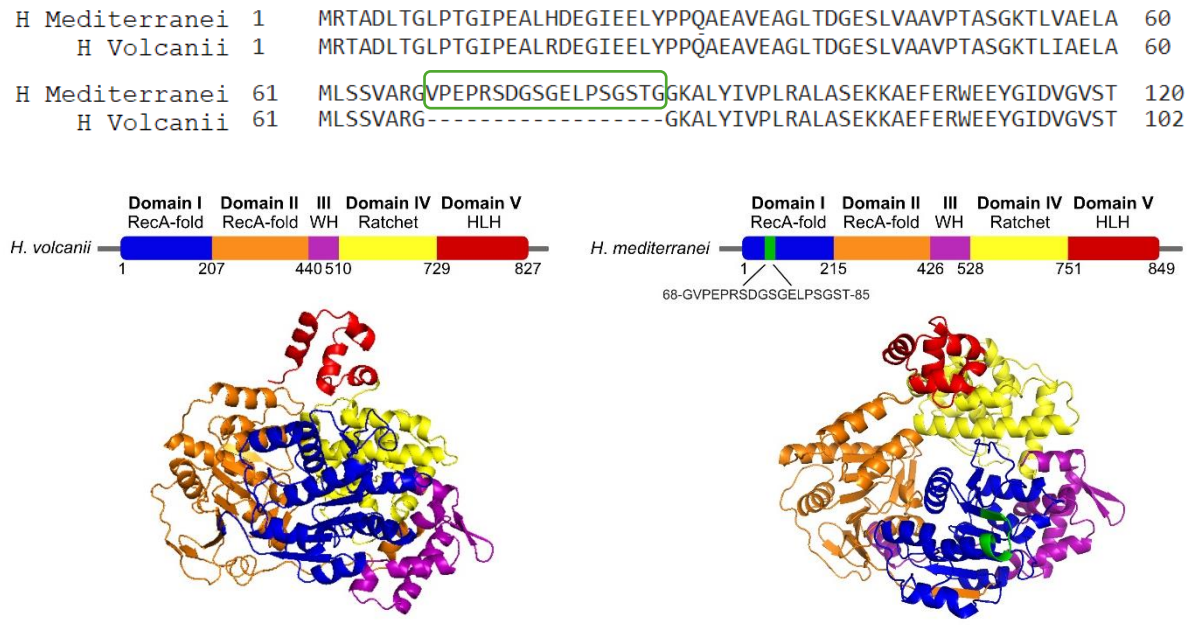


Figure 21. Comparison of structures of H. volcanii hel308 *(left) and* H. mediterranei hel308 *(right). Domains are colour-coded; the 18 amino acid sequence unique to* H. mediterranei *is shown in green. Figure courtesy of Dattani, unpublished.*

Overexpression of *hel308* is known to be toxic (Thorsten Allers, personal communication). In addition, as originless *H. volcanii* strains are dependent on *radA* for replication, and are thought to be replicating via recombination, it is possible that Hel308 could interfere with this process by unwinding D- and R-loops, leaving the cells unable to divide due to ineffective DNA replication. It should also be considered that the observed effect may not be unique to the *H. mediterranei* Hel308, but may be due to increased dose of Hel308 within the cell; after all, the strains still possessed their own chromosomal copy of *hel308*, in addition to the *H. mediterranei* copy. It is

possible that OIR was inhibited by the increased dose and/or activity of Hel308, rather than *H. mediterranei* Hel308 specifically.

3.2 Aims and Objectives

In order to determine whether the observed effect was indeed due to the Hel308 gene present in the genome library fragment, and whether this was an effect unique to this species, the same strain and vector of the screen was used in a verification experiment. Instead of a large *H. mediterranei* genome fragment, the *hel308* gene from this organism would be inserted into the plasmid used in the screen described above, and the effects of this gene specifically would be observed. To provide further insight, additional plasmids bearing *H. volcanii hel308* would also be produced, to ensure that the effect was unique to the *H. mediterranei hel308*. A final plasmid containing the *H. mediterranei hel308* gene with a small truncation (G68-T85del) would also be produced, in case this elucidated the cause of the effect. Following confirmation of the screen results, further examination could be made of the properties of Hel308 that allow or prevent OIR, such as truncated genes and point mutants.

The screen would also be subject to further verification through inducing expression of *H. mediterranei* or *H. volcanii* Hel308 in originless strains of *H. volcanii*; if the genes interfere with origin-independent replication, then the cells would die upon their induction. This reduces the potential for any interference with control of the inducible origin itself; there is no possibility that an introduced factor could lead to expression of *orc10*, or interact with the origin. The survival of the cells is therefore at all times directly dependent on their continued ability to undergo origin-independent replication.

To these ends, the objectives of this chapter were therefore to:

- Insert the *hel308* gene from both species into the vector used in the screen.
- Transform these plasmids into the specific *H. volcanii* strain used in the screen.
- Explore the viability of these strains in the presence and absence of tryptophan.

And as an additional check in a fully origin-deleted background:

- Produce plasmids which allow inducible expression of *hel308* from both species.
- Transform these plasmids into origin-less *H. volcanii* strains.
- Explore the viability of these strains in the presence and absence of tryptophan.

3.3 Confirmation of screen results

3.3.1 Plasmid construction

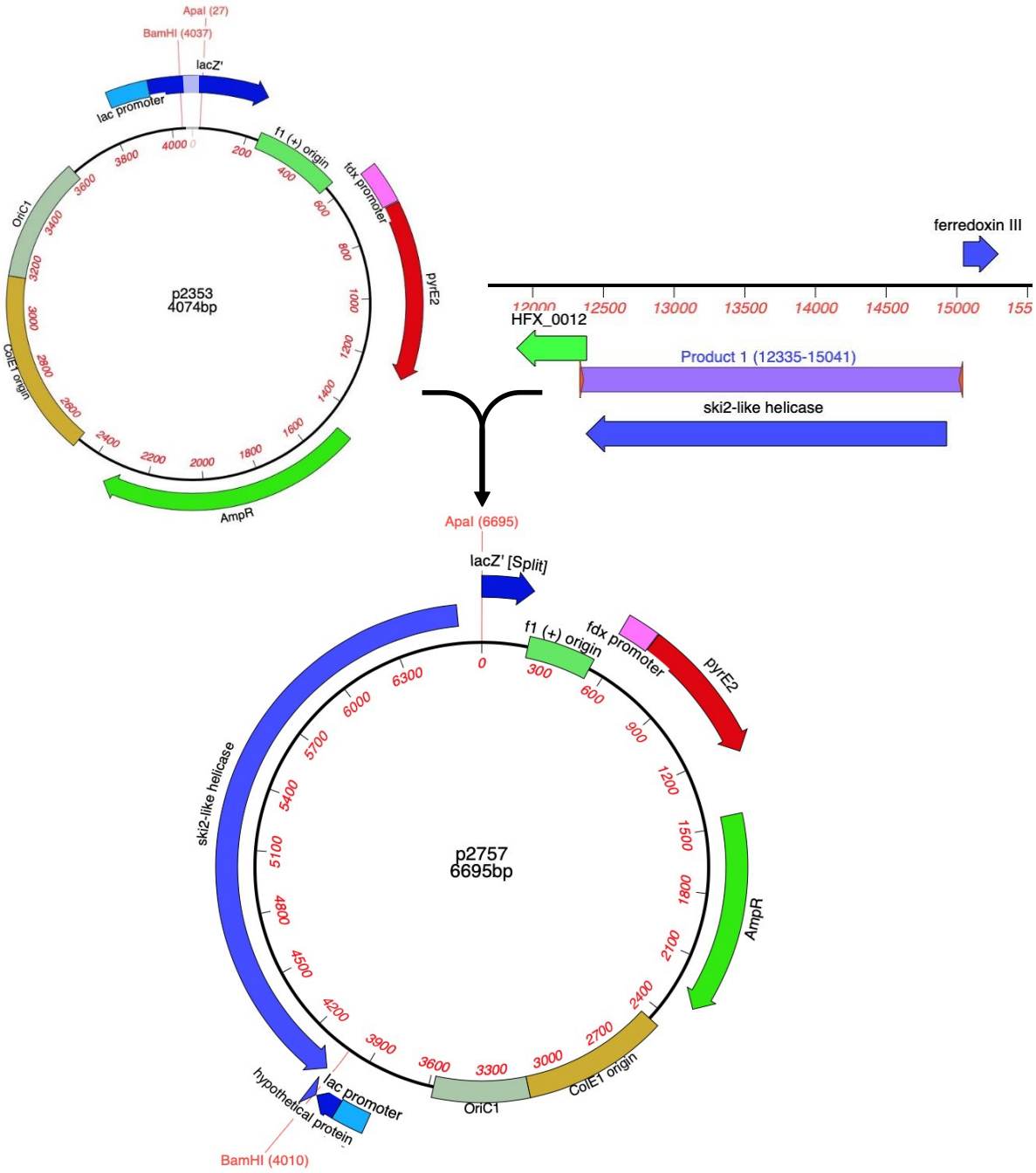


Figure 22. Creation of p2757 - plasmid bearing H. mediterranei hel308. This gene is annotated 'ski2-like helicase'.

The *hel308* gene from *H. mediterranei* was cloned from the genome of H824 (*ΔpyrE2*) using primers o2417 and o2379, as shown in Figure 22 (above). The forward of these primers binds around 100bp upstream of the start codon, in an attempt to ensure inclusion of the promoter region and encourage expression at near-physiological levels. Insertion at the ClaI site (as originally intended) was not possible due to the presence of this restriction site within the gene. The PCR product and p2353 vector were digested with ApaI and BamHI and ligated to produce p2757. Insertion of the gene was confirmed by diagnostic digest and sequencing.

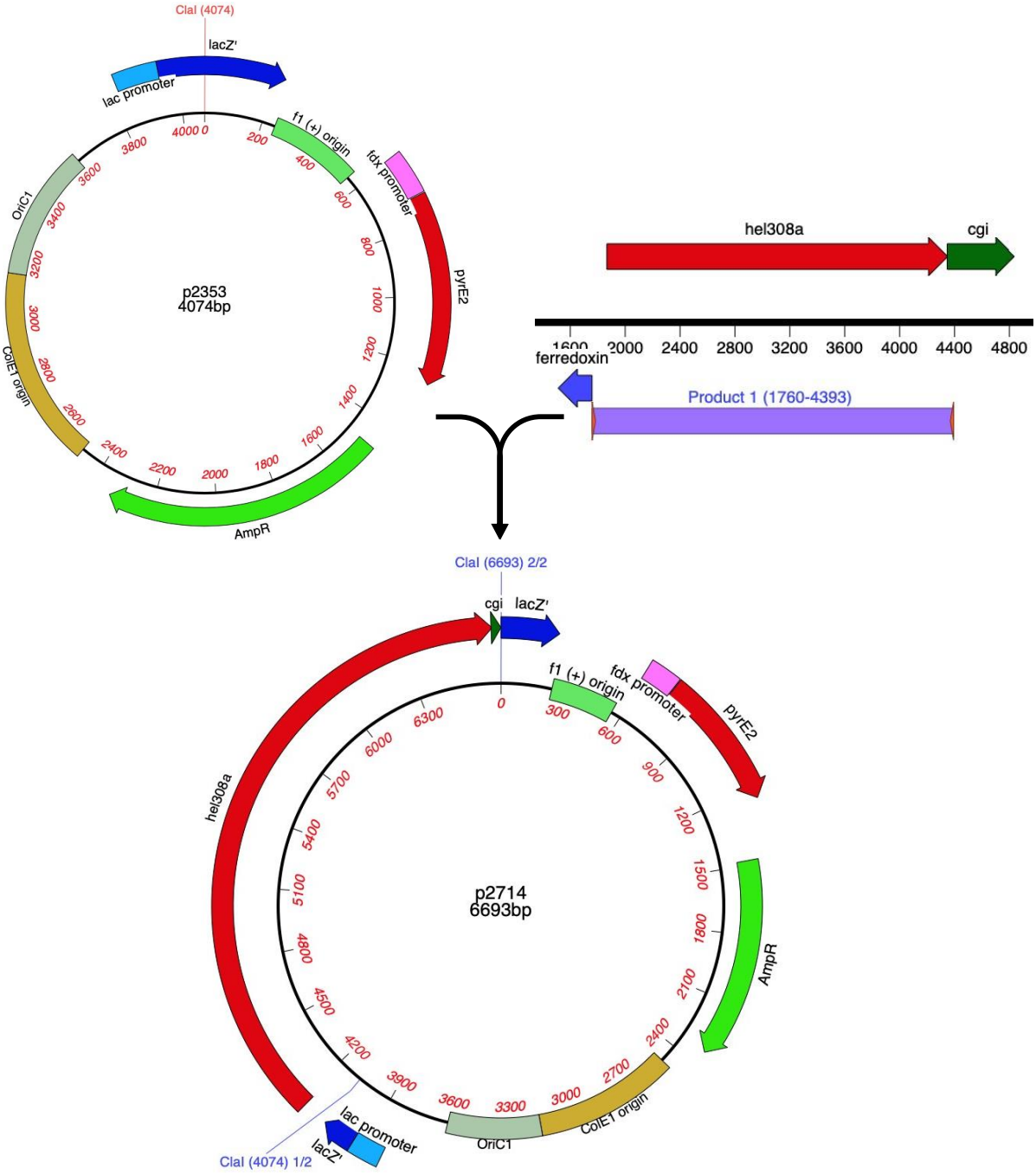


Figure 23. Creation of p2714 - plasmid bearing H. volcanii hel308

The *hel308* gene from *H. volcanii* was cloned from pTA1316, itself a genomic clone of this gene, using primers o2370 and o2371, which both introduce ClaI sites to the product. The PCR product and p2353 vector were digested with ClaI and ligated to produce p2714. Insertion of the gene (and its orientation) was confirmed by diagnostic digest and sequencing. It should be noted that, as only a single restriction enzyme was used to cut the vector, the gene was able to insert in the opposite orientation to the *H. mediterranei hel308* equivalent (above). However, as the native promoter is included, this should not affect gene expression levels. The lacZ promoter is present on the plasmid to allow blue/white selection in *E. coli*; it is not active in *Haloferax*.

3.3.2 Strain construction

Table 3-1 Strains used in this chapter. Note that the parent strains, H5074, and H5107, are Δmrr, meaning that plasmids lacking dam methylation do not need to be produced.

Strain	Genotype	Notes
H5107	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *ΔoriC1::[oripHV1 p.tnaA M3::orc10 LeuB+]*	Produced by A. Dattani. H5074, with oripHV1 and inducible orc10 inserted at original oriC1 position.
H5526	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *ΔoriC1::[oripHV1 p.tnaA M3::orc10 LeuB+]* *{hel308+ pyrE2+}*	Produced for this study. H5017 with episomal *H. volcanii hel308*
H5527	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *ΔoriC1::[oripHV1 p.tnaA M3::orc10 LeuB+]* *{Hmedhel308+ pyrE2+}*	Produced for this study. H5017 with episomal *H. mediterranei hel308*
H77	ΔtrpA	Produced by T. Allers. Used as a control in tryptophan gradient plates.

3.3.3 Confirmation of screen results using original strain and vector

The plasmids and strains were constructed as per the Methods chapter, and were used to perform a tryptophan gradient assay. If the genes present on the plasmid were able to inhibit origin-independent replication, the strains would not grow on the side of the plate without tryptophan, as the origins would be inactive in this environment. Care was taken to maintain tryptophan exposure of the strains at all culturing steps prior to plating. The results are shown in Figure 24, below.

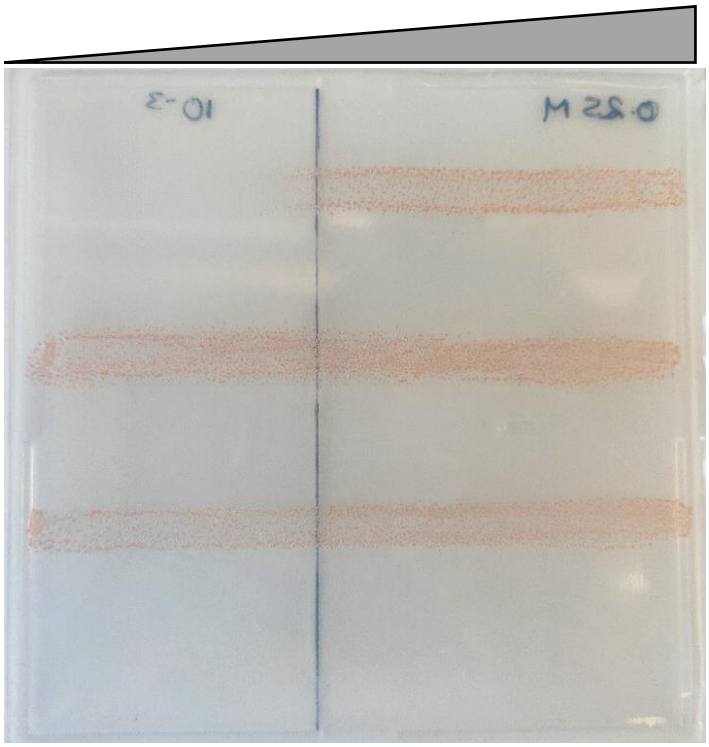


Figure 24. Tryptophan gradient plates, with tryptophan gradient increasing from left to right across each plate to a final concentration of 0.25mM (left) and 1mM (right). Narrow edge of the tryptophan media segment is indicated by the black line. Plates show the growth of H77 (top), H5526 (middle) and H5527 (bottom) across the gradient after four days in a 45°C incubator. H77 is a tryptophan auxotroph, confirming the presence of the gradient.

Unexpectedly, H5526 (containing the *H. volcanii hel308* gene) and H5527 (containing the *H. mediterranei hel308* gene) were able to grow all the way across the plate, even on the extreme left side which contains no tryptophan. Persistence of the plasmid in the strains was ensured by excluding uracil from the media; cells that have lost the plasmid would revert to uracil auxotrophy and be unable to grow. Presence of the *hel308* gene was confirmed by sequencing one colony from each strain, picked from the leftmost side of the plate, where the cells were thought to be undergoing origin-independent replication. This also confirmed that the plasmid had not integrated onto the chromosome, thus inadvertently supplying an oriC1 origin.

As the strain background and plasmid backbone are identical to those used in the screen, this suggests that the effect observed in the screen was not due to the presence of the *H. mediterranei hel308* gene, or increased Hel308 production or activity within the cell. In short, the presence of *H. mediterranei hel308* gene is insufficient to prevent origin independent replication in *H. volcanii*.

3.4 Confirmation of screen result using inducible hel308

3.4.1 Plasmid construction

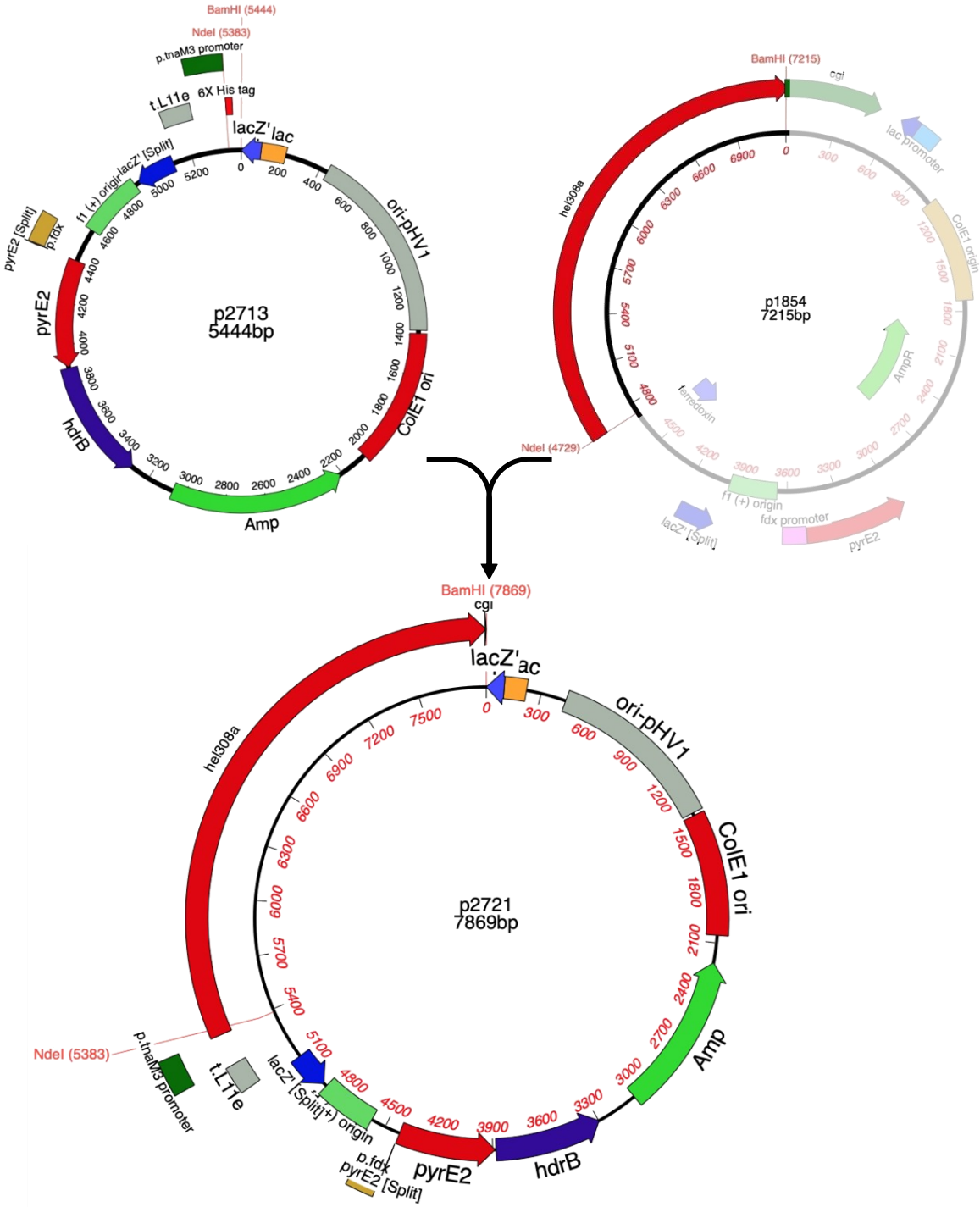


Figure 25: creation of inducible H. volcanii *hel308 plasmid.*

The inducible *H. volcanii hel308* plasmid was produced for Chapter 4 by Msci student Olivia Wood, under supervision. Vector p2713 was selected for that project for its combination of *p.tna*M3 promoter, which reduces expression levels by around 50% compared to *p.tnaA* (Braun et al., 2019). In addition, the pHV1 origin ensures low copy number of the plasmid within the *Haloferax* cell. These characteristics were chosen to ensure lower expression upon induction – Hel308 protein is naturally expressed at low levels, and is known to be toxic at high levels. The His-tag (used for protein purification) is not needed for this application, and is removed during gene insertion between

NdeI/BamHI restriction sites. t.L11e indicates a transcription terminator to prevent through-transcription from other parts of the plasmid (Shimmin & Dennis, 1996).

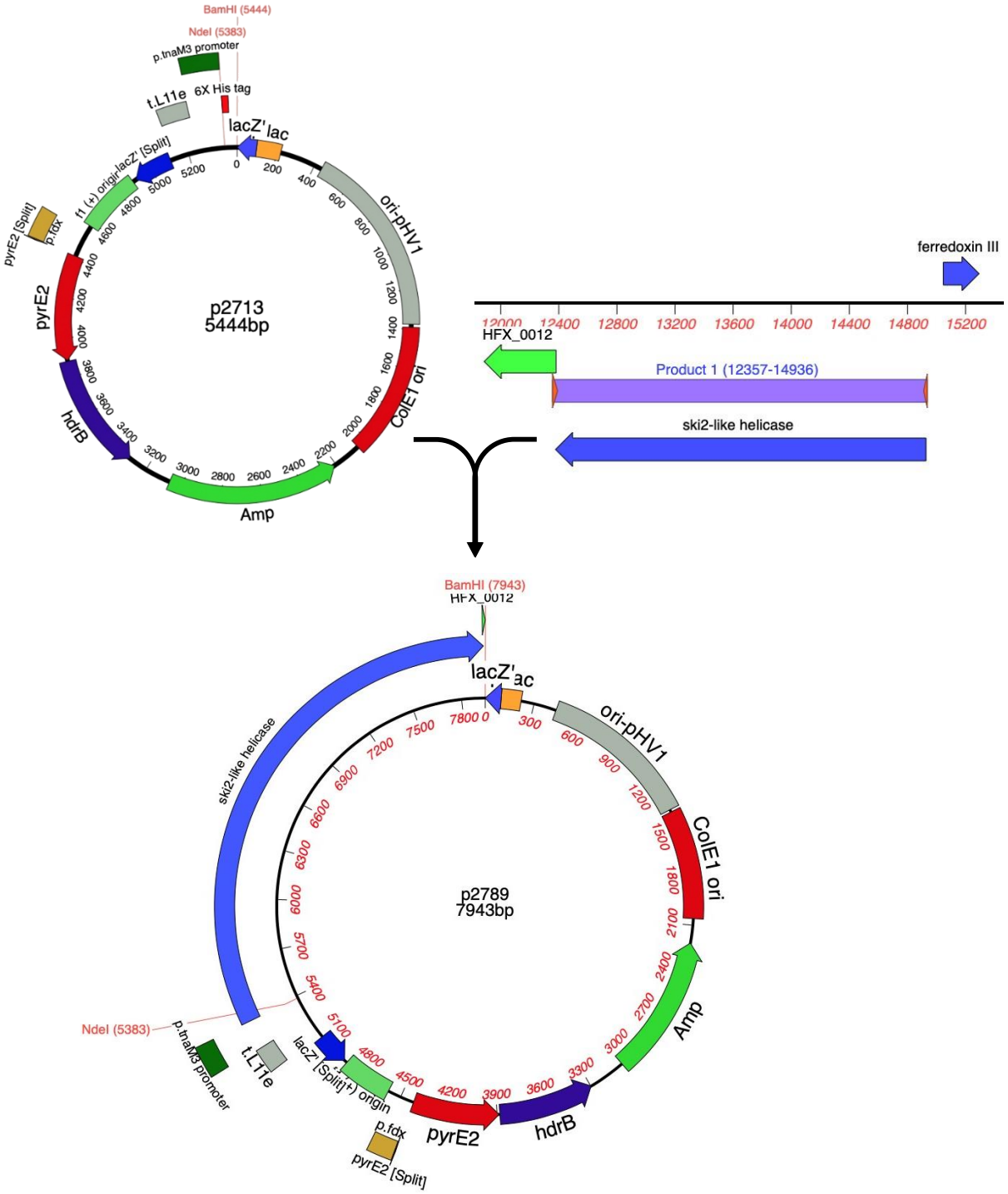


Figure 26 Creation of inducible H. mediterranei *hel308 plasmid.*

Plasmid p2789 was also made from the p2713 vector. The *hel308* gene was cloned from *H. mediterranei* strain H826, using the primers o2274 and o2464, introducing the NdeI and BamHI sites. Insertion of the gene removes the His-tag, which is not needed for this application.

3.4.2 Strain Construction

Strains used in this section are summarised in the table below.

Table 3-2 Strains used in this section

Strain	Genotype	Notes
H1391	*ΔpyrE2* *Δhel308*	Produced by T. Allers Strain with chromosomal hel308 deleted. Used to confirm expression by inducible plasmids.
H5074	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3*	Produced by A. Dattani. *H. volcanii* strain lacking all origins, and with only a single chromosome.
H5613	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *{p.tnaA::hel308+ pyrE2+ hdrB+}*	Produced for this study. H5074 with inducible *H. volcanii hel308*
H5614	*ΔpyrE2* *Δmrr* *ΔoriC1* *ΔoriC2* *ΔoriC3* *ΔoripHV3* *Δori-pHV4-2* *ΔpHV1* *ΔleuB::[pHV3]* *Δ3'-5'LeuB::pHV3* *{p.tnaA::Hmed hel308+ pyrE2+ hdrB+}*	Produced for this study. H5074 with inducible *H. mediterranei hel308*
H5608	*ΔpyrE2* *Δhel308* *{p.tnaA::hel308+ pyrE2+ hdrb+}*	Produced for this study – expression assay for inducible plasmids
H5609	*ΔpyrE2* *Δhel308* *{p.tnaA::Hmedhel308+ pyrE2+ hdrb+}*	Produced for this study – expression assay for inducible plasmids
H5533	*ΔpyrE2* *Δhel308* *{pyrE2+}*	Produced for this study – empty vector control for expression assay

3.4.3 Assay for inducible *hel308* plasmids

In order to confirm that the *hel308* homologues were being successfully expressed by p2721 and p2789, these plasmids (and an empty vector control; p357) were transformed into a *Δhel308* strain (H1391). These strains are summarised below:

Table 3-3 Strains used in this section

Strain	**Genotype**	**Notes**
H5533	*ΔpyrE2* *Δhel308* *{pyrE2+}*	Produced for this study. Empty vector control
H5608	*{p.tnaA::hel308+ pyrE2+ hdrb+}*	Produced for this study. Inducible *H. mediterranei hel308*
H5609	*{p.tnaA::Hmedhel308+ pyrE2+ hdrb+}*	Produced for this study. Inducible *H. volcanii hel308*

As deletion of *hel308* confers a growth defect on the strains, successful expression of the inducible *hel308* alleles can be easily assessed by measuring growth rate in the presence and absence of tryptophan, to observe complementation of the growth defect. As an additional benefit of this assay, it can also be determined whether *H. mediterranei hel308* is also capable of complementing this defect, or whether it lacks some relevant property of *H. volcanii* Hel308 in this regard.

As it is difficult to quantify growth rate from observing colony size, this assay was instead carried out via growth assay in a spectrophotometer. Strains H5533, H5608 and H5609 were grown in plain Cas broth overnight, and were then used to inoculate either plain Cas broth, or Cas plus 0.25mM tryptophan, and left to grow for 72 hours, with measurements taken every 15 minutes. 0.25mM of tryptophan has previously been shown to be sufficient to induce expression through *p.tnaA*, as shown in Large et al. (2007).The results of this assay are shown in Figure 27, below, with ten biological replicates for each strain and condition combination.

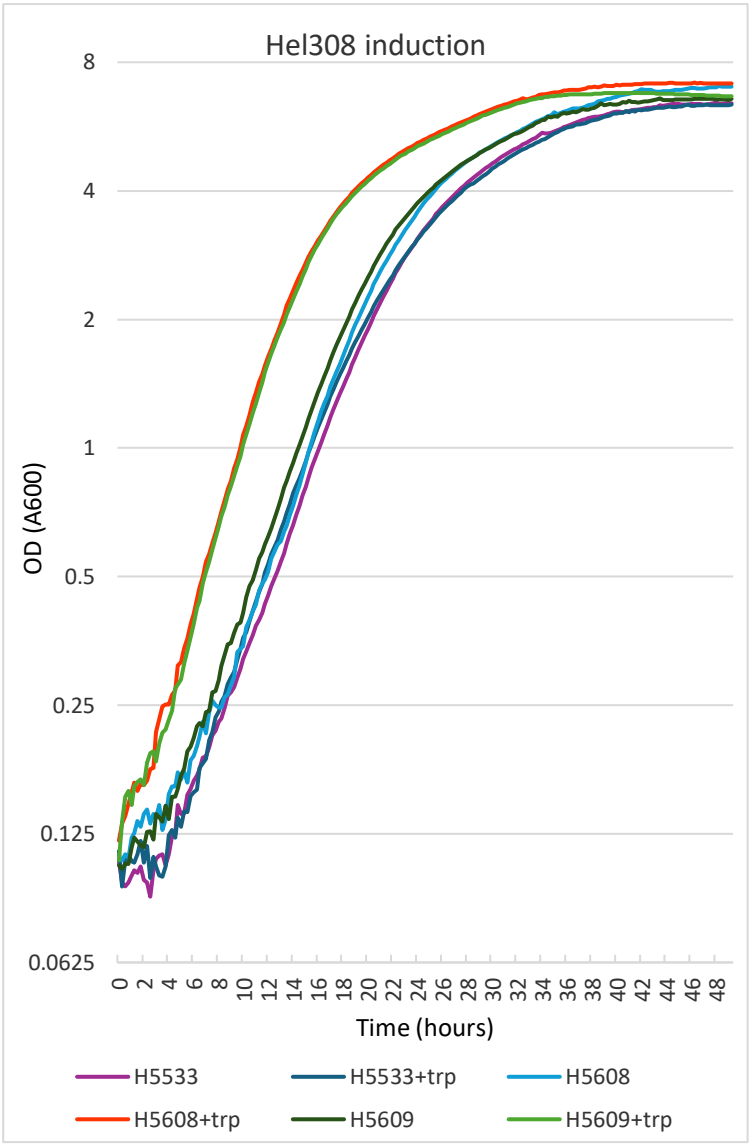


Figure 27 Inducible hel308 *test. "+trp" indicates cultures growing in Cas broth with 0.25mM of tryptophan added.*

As can be seen from the above graph, the induction of *hel308* under 0.25mM tryptophan is sufficient to complement the growth defect observed in *Δhel308* strains. This not only shows that expression levels are sufficiently high, despite the *p.tna*M3 promoter and low-copy-number origin, but also shows that the *H. mediterranei hel308* homologue is sufficient to complement deletion of this gene in *H. volcanii*. The plasmids are therefore fit for purpose.

3.4.4 Confirmation of screen using inducible *hel308*

Expression of the gene of interest in the origin-deleted strain allows certainty that the cells are still undergoing origin-independent replication. The production of the plasmids and strains for this experiment was begun before the results from the test using the original screen strain were attained. A tryptophan assay was performed as per the Methods chapter. The results are shown in Figure 28, below.

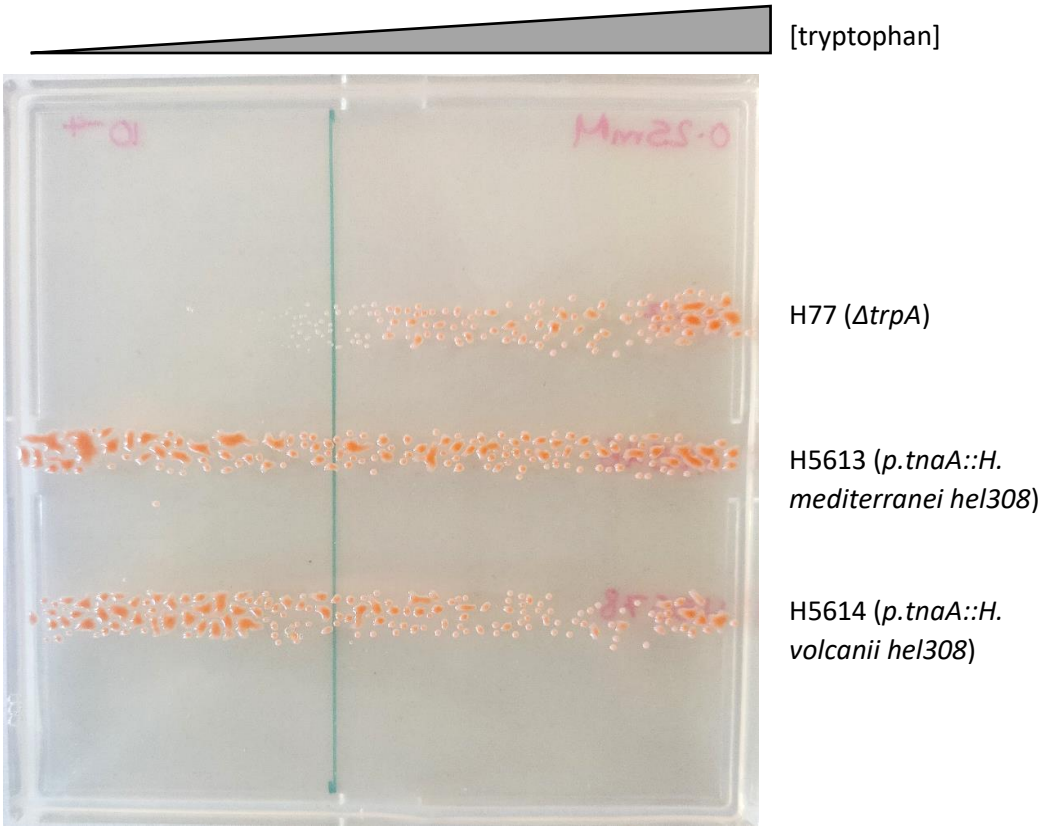


Figure 28. Tryptophan gradient plates, with tryptophan gradient increasing from left to right across each plate to a final concentration of 0.25mM. Narrow edge of the tryptophan media segment is indicated by the green line. Plates show the growth of H77 (top), H5613 (middle) and H5614 (bottom) across the gradient after four days in a 45°C incubator. H77 is a tryptophan auxotroph, confirming the presence of the gradient.

Inducing the expression of the *hel308* homologues did not affect the viability of the strains, which grew all the way across the gradient. This was the same in both low concentrations of tryptophan, and high concentrations (data not shown). While the presence of the plasmids in the strains is effectively confirmed by their continued uracil autotrophy, it was possible that the plasmids had spontaneously lost part of the gene (as reported in Lever (2020)). In Lever's work, attempts to characterise a specific point mutant were hampered by the spontaneous loss of the plasmids bearing the gene of interest, either in whole or in part. To ensure that an equivalent phenomenon was not masking the effects of the *hel308* gene in this experiment, samples were taken from the high-tryptophan sides of the plates, and the presence of both the plasmid and the *hel308* gene were confirmed by sequencing.

3.5 Further exploration and discussion of results

Following the unexpected results above, the original strains identified in the screen were re-streaked onto YPC plates from the glycerol stocks stored in a -80°C freezer. The plasmids from these strains were extracted via a phenol-chloroform procedure (described in the methods chapter, above) and transformed into *E. coli*. This step is necessary as the yield and quality of plasmids extracted from *Haloferax* are typically quite low (T. Allers, personal communication). Midipreps of the plasmids were produced, to ensure a high concentration of good-quality DNA. These plasmids were then sent for sequencing. However, coherent sequences could not be recovered from them.

It was considered that partial loss of the plasmid could have occurred during culturing prior to freezing of the glycerol stocks. Individual colonies from the streaked-out stocks (which had been streaked out on non-selective YPC media) were patch-plated onto Cas-based media, both with and without tryptophan. Not all of the colonies grew under these conditions; however, there were no differences found between the two replica plates (examples shown in Figure 29, below). All colonies that failed to grow on the plain Cas plates also failed to grow on the Cas+trp plates. Given the genotype of the parent strain, it is likely that these colonies failed to grow due to a reversion to uracil auxotrophy, suggesting that they had lost the *pyrE2*-marked plasmid altogether. Even in those that still possessed the plasmid, whatever factor had previously produced their tryptophan-dependent phenotype was seemingly no longer present.

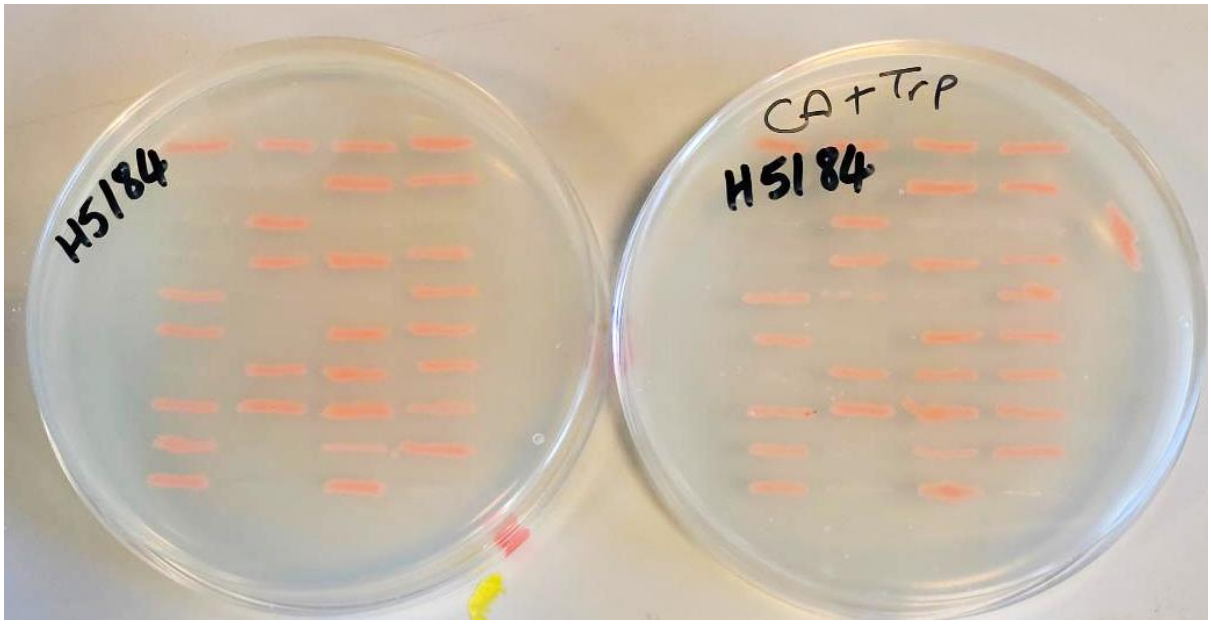


Figure 29. H5184 (candidate bearing H. mediterranei hel308) identified in A. Dattani's screen) following restreaking onto Cas and Cas+trp plates. Growth is identical on both plates.

This is unfortunate, as the tryptophan-dependent growth for several of the screen candidates (including H5184) had previously been a reliable phenotype, and had been demonstrated by both re-streaking onto different media, and through tryptophan gradient assays, as show in Figure 30, below.

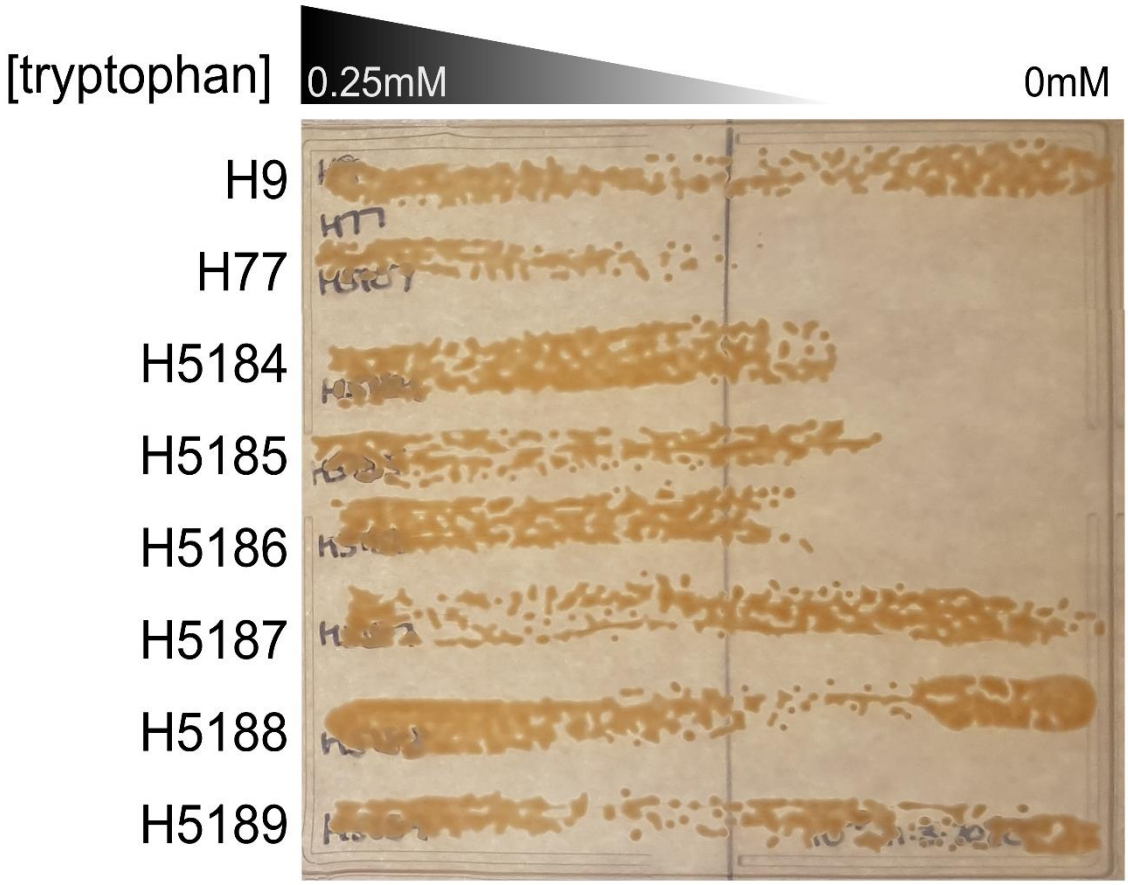


*Figure 30. Tryptophan gradient plate showing the reliance of strains H5184-6 on tryptophan for their growth. H77 (*ΔtrpA*) included as a control to visualise the tryptophan gradient. H5187-9 were control colonies from the screen which did not exhibit tryptophan-dependent growth. From A Dattani (personal correspondence).*

It seems likely that the plasmid producing the effect may have been spontaneously lost while the strains were being cultured before freezing. *Haloferax* is prone to genome rearrangements (Ausiannikava et al., 2018), and plasmids are often not perfectly stable in terms of persistence in the strain. However, the plasmids in question contained the *OriC1* origin, which has been shown to confer good plasmid persistence (Norais et al., 2007).

However, there may have been a selective pressure towards loss of the plasmid. If the plasmid identified in the screen was capable of causing cell death in the absence of origin activity, it may also have been causing some defect to growth or cell function even while the origin was active. This would grant a comparative growth advantage to any cells in the culture that had lost the plasmid, effectively providing selective pressure for its own loss. This could explain why the plasmid was not found in the cells recovered from the frozen stocks.

It is also possible that, when the *H. mediterranei* genome library was transformed into H5107, some of the transformed cells may have received multiple different plasmid copies at once, or plasmids containing multiple genome fragments ligated together. While *H. mediterranei hel308* may have been involved in causing the tryptophan-dependent phenotype observed, the effect could have been due to multiple genome fragments acting in combination. As all of the relevant plasmids would have borne the same marker genes (only differing in the inserted genome library segment), the loss of some of the variants from the strain would not have been detectable without sequencing.

3.6 Future exploration

While the specific plasmids that caused the tryptophan-dependent phenotype in the screening experiment appear to have been irrecoverably lost, the genome library that was used to create them still exists, as does the parent strain. It would be possible to repeat the screen to see if the same *H. mediterranei* genome sections once again make their presence known. However, the investment of time and effort in this undertaking should not be underestimated. During the original screen, six tryptophan-dependent colonies were identified following the replica plating of 5000 candidate colonies. As there was no repetition between these tryptophan-dependent colonies (each genome section of interest was identified only once), it can be assumed that these are very rare among the total population. There is therefore no guarantee that the same genome fragments would be identified during a repeat of the screen using similar numbers of candidates. If we assume the "true" frequency of the hel308-bearing plasmid in the genome library to be 1 in 5000, the chances of it being identified again in a screening of a further 5000 colonies is 63.21% ($1-\left(\frac{4999}{5000}\right)^{5000}$). This means that, in a repeat of the screen, there is a 36.78% chance that much time-intensive work could be invested to no gain.

3.6.1 Alternate avenues of exploration

As mentioned previously, the inability of *H. mediterranei* to undergo origin-independent replication may be due to its lower rate of recombination compared to *H. volcanii* (Dattani, Sharon, et al., 2022). These lower levels may not provide sufficient recombination events for effective genome replication. Future work could explore artificially increasing the recombination rate of *H. mediterranei* cells, to determine whether this allows deletion of all four *H. mediterranei* chromosomal origins. For example, factors that increase recombination could be introduced, such as the *H. volcanii hel308-F316A* point mutant (Lever, 2020).

3.6.2 Proposal for an alternative screen

Another method to identify OIR-inhibiting factors in *H. mediterranei* involves hybridisation of cells between the two species. *H. volcanii* and *H. mediterranei* are able to fuse to generate hybrid cells; within these cells, recombination events can result in chromosomes containing sections from each species (Naor et al., 2012). Naor et al determined that these chromosomal exchanges were typically between 300kb-500kb in length. While hybrid cells are thought to be initially heterodiploid, containing a variety of different chromosomes, lineages tend towards a more homozygous state over generations (Dattani et al., 2022).

Mating of a *H. volcanii* strain in which all chromosomal origins have been deleted, with a *H. mediterranei* strain in which all but one origin has been deleted (and the final origin marked with an adjacent marker gene such as *pyrE2*), would result in fused cells containing both genomes. As these fused cells grow, a combination of recombination events between homologous sequences and random chromosome segregation during cell division would result in a variety of offspring, hopefully with a variety of novel chromosomes containing alleles from both parents, as shown in a simplified form in Figure 31.

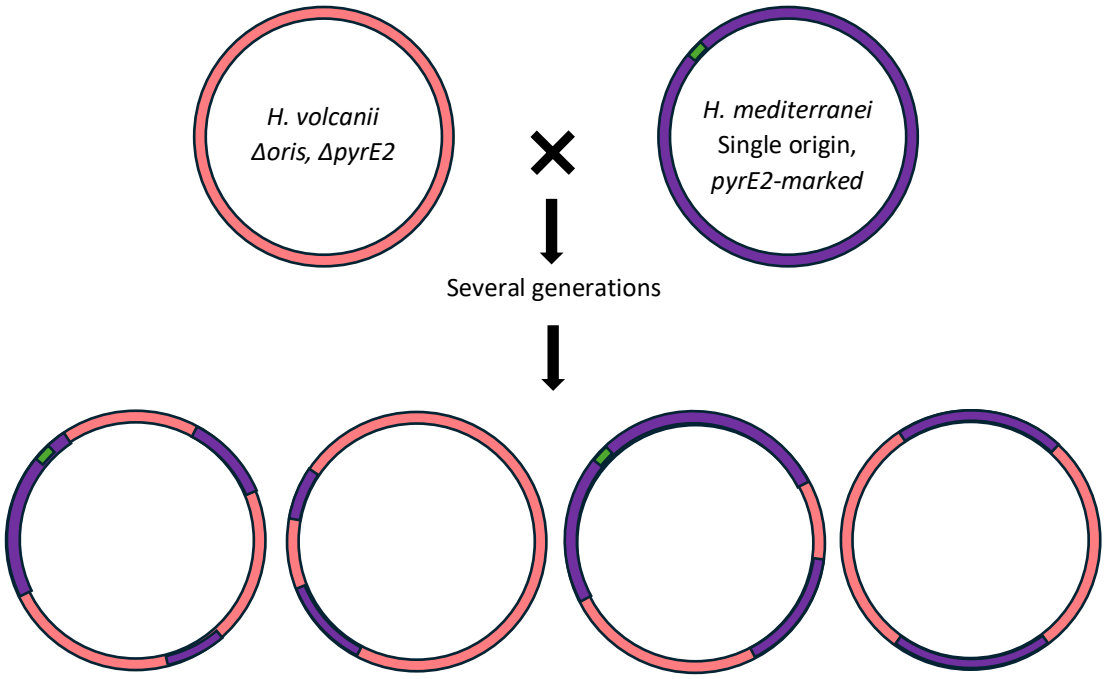


Figure 31. Hybridisation of origin-less H. volcanii *(red) strain with single-origin* H. mediterranei *strain (purple). Position of the remaining* H. mediterranei *origin, bearing a* pyrE2 *marker gene, is indicated in green (position arbitrary). Offspring bear chromosomes with a mixture of alleles from both parents; only some inherit the single origin.*

Plating the hybrids on plates containing 5'FOA and uracil would result in isolation of strains that do not contain the only remaining origin. Sequencing a selection of these hybrids would allow some mapping of where the OIR-inhibiting factors of *H. mediterranei* can be found in the genome. Any *H. mediterranei* portions of the genome that can be found in cells without origins are presumably not those that carry factors that interfere with OIR. With enough hybrids sequenced, hopefully some candidate regions of the *H. mediterranei* genome can be identified for further study; those which are consistently not seen in originless hybrids. A simplified example is shown in Figure 32.

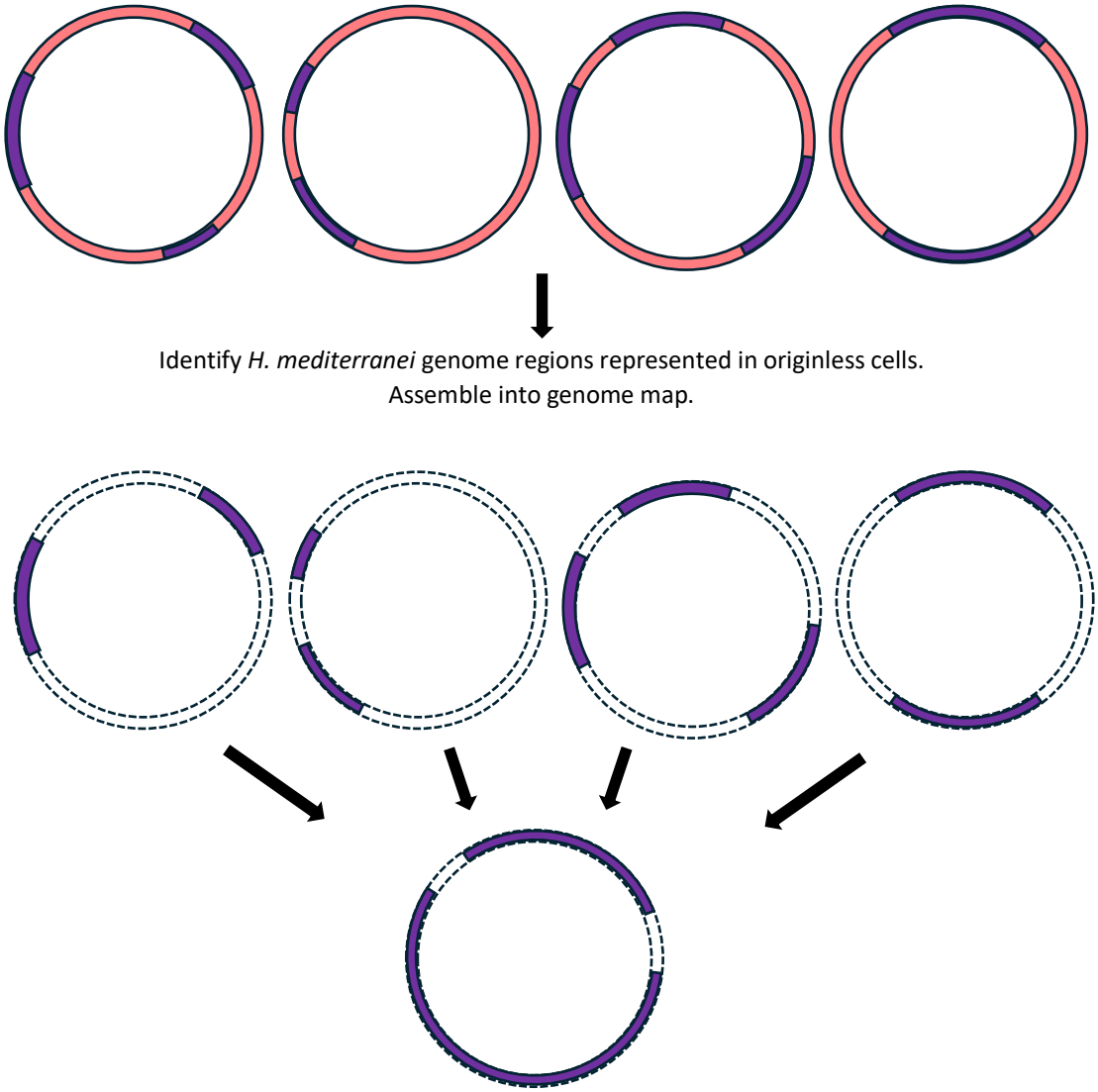


Figure 32. Compiling all represented H. mediterranei *genome fragments from this example reveals two regions that are not seen in originless cells; one region in the top left (the position of the origin) and one region on the right, which may contain a factor incompatible with origin-independent replication.*

To further develop this concept, and to ensure that all offspring are recombinant, two different auxotrophic markers could be included, one in each parent, and only cells bearing both markers selected for subsequent steps. However, this would introduce some bias into the screen, as the regions of the chromosomes bearing the marker genes would be of known parentage in all cases, as well as at least some of the proximal genome segments.

Were this method to fail to identify any obvious candidate regions, it could be that multiple factors acting in combination are responsible for preventing OIR in *H. mediterranei*. It could also be that the position of the *pyrE2*-marked origin happens to be close to some of the OIR-inhibitory genome sections. It has previously been reported that all *H. mediterranei* origins are non-essential (Yang et al., 2015), and that deletion of any three is thought to be possible. Therefore, to alleviate the positional bias, the experiment could be repeated, with a different *H. mediterranei* origin selected to be the 'final' marked origin.

As a final consideration; it is possible that *H. mediterranei*'s inability to undergo origin-independent replication is not due to a preventative factor in *H. mediterranei*, but rather the lack of a permissive or enabling factor present in *H. volcanii*. If this is the case, the screening tool suggested above could

alternatively be used to identify regions of the *H. volcanii* genome that are consistently seen in originless hybrids; these may contain essential factors in the replication of the genome through origin-independent methods.

3.7 Conclusion

The Hel308 proteins of *H. volcanii* and *H. mediterranei* differ slightly, much like their host organisms. It was previously claimed that *H. mediterranei hel308* could inhibit origin-independent replication in a *H. volcanii* model. However, expression of *H. mediterranei hel308* in the previously-used model could not reproduce the same effect. Expression of inducible *hel308* also could not inhibit the growth of originless *H. volcanii* cells.

The factors that originally induced the phenotype of interest in previous work have not been identified, but it is concluded that expression of *H. mediterranei hel308* is insufficient to produce the effect reported. It is, however, sufficient to complement the growth defect observed in *Δhel308 H. volcanii* strains.

4. Potential lethality of *hel308-K53A*

4.1 Background

4.1.1 Previous work on Hel308 RecA point mutants

The structures and functions of *H vol* Hel308 have been previously investigated via generation of point mutations.

An ATPase-null mutant was created by disrupting the DEAD/H motif in Domain 1 (Lever, 2020). In this motif, the negatively charged aspartic acid-145 is necessary for interaction with positively charged Mg^{2+}, which is required for ATP hydrolysis. Mutants without this negatively charged residue should therefore be unable to catalyse ATP hydrolysis, and consequently unable to translocate along the bound DNA, abolishing helicase activity. Strains carrying the *hel308-D145N* mutation (underlined DEAD/H) as a gene replacement exhibit a 50% decrease in recombination levels, meaning that they are not equivalent to *Δhel308* mutants (which demonstrate a recombination rate 5.5 times higher than wild type) (Lever, 2020). This suggests that *hel308* has both some ATP-dependent and some ATP-independent roles in recombination.

However, the response to DNA damage in hel308-D145N strains (administered by ultraviolet irradiation or MMC administration) is equivalent to that of the *Δhel308* strain (Lever, 2020). This implies that ATPase activity is essential to Hel308's role in DNA repair, but that some non-ATPase dependent activity is relevant to its role in recombination.

A corresponding disruption of the Walker A motif was also attempted. Lysine-53 is located in Domain 1, within the RecA fold. A *hel308-K53A* point mutant was produced, in which the small, uncharged alanine should hinder breakdown of ATP by this domain (Lever, 2020). However, repeated transformation attempts failed to yield mutants, and in instances where the transformation appeared successful, further examination revealed that the relevant part of the gene had been spontaneously lost (Lever, 2020). It was hypothesised that *hel308-K53A* is in fact a lethal mutant, although it is unclear why this should be the case. In theory, disrupting this other part of the ATP-binding domain should yield the same phenotype as the other ATPase-null mutant described above. However, it is possible that a version of the enzyme unable to bind ATP might behave differently to a version of the enzyme that is able to bind ATP, but not catalyse it.

Point mutants equivalent to K53A have been produced in other homologues. In *S. solfataricus,* the equivalent point mutant K52A has been studied, although this protein was produced in *E. coli* rather than the organism of origin (Richards et al., 2008). This enzyme was found to have DNA binding affinities equivalent to wild-type, but was unable to separate DNA strands. The *Methanothermobacter thermautotrophicus hel308* point mutant K51L was also produced in *E. coli*, and was found to be inactive as a helicase (Guy & Bolt, 2005).

Walker A and B mutants were also produced for the *Pyrococcus furiosus* homologue (K52A and D145A respectively (Fujikane et al., 2006)). While these mutants were found to be ATPase-null, they were still able to partially complement the growth defect of *E. coli* cells carrying a *recQ* mutation, suggesting both ATP-dependent and ATP-independent roles. Expression of both of these proteins in *E. coli* hosts shows that they are not lethal to this bacterium.

One potential flaw introduced by characterising these proteins expressed in heterologous hosts is that archaea are capable of a wide range of post-translational modifications, including glycosylation, phosphorylation, methylation, and sampylation; tagging proteins with SAMP (small archeal modifier protein) (Darwin & Hofmann, 2010). These are reviewed in Gong et al. (2020). However, it is possible

that the form of these proteins expressed in the native host would bear modifications which alter its activity, which would not be replicated in the heterologous host. It is not currently known whether *H. volcanii* Hel308 is affected by post-translational modification.

4.2 Aims and Objectives

It was thought that introduction of *H. volcanii hel308-K53A* under the control of an inducible promoter would allow transformation of this gene into live cells, and control its expression. This would allow its lethality to be verified, and perhaps allow exploration of its mechanism.

The *p.tnaA* promoter allows very tight repression in the absence of tryptophan (Large et al., 2007), hopefully allowing a lethal gene to be present in the cell without affecting its viability. The effects upon expression could then be examined. Use of this promoter has allowed exploration of scenarios which would be otherwise lethal to the cell. For example, see Liao et al. (2021), in which an essential gene was placed under tryptophan-inducible control. The cells could be first deprived of the encoded protein, exploring the effects of its absence on the cell, which would not be possible by deletion. Expression could then be induced, resupplying the cells with the encoded protein. The effects of both expression and repression could then be examined. It was thought that the effects of *hel308-K53A* could perhaps be explored in a similar manner.

However, the *p.tnaA* promoter does not inhibit expression effectively in *E. coli*. As this gene was also reported to be lethal in this bacterium, the usual cloning protocols would be adapted to forego steps involving *E. coli*, and constructed plasmids would be transformed straight into *Haloferax*. The objectives are therefore to:

- Produce a vector for the inducible expression of various *hel308* point mutants
- Transform these plasmids into *H. volcanii* strains
- Explore the viability of these strains in the presence and absence of tryptophan

4.3 Hel308-K53A lethality test

4.3.1 Plasmids produced for this experiment

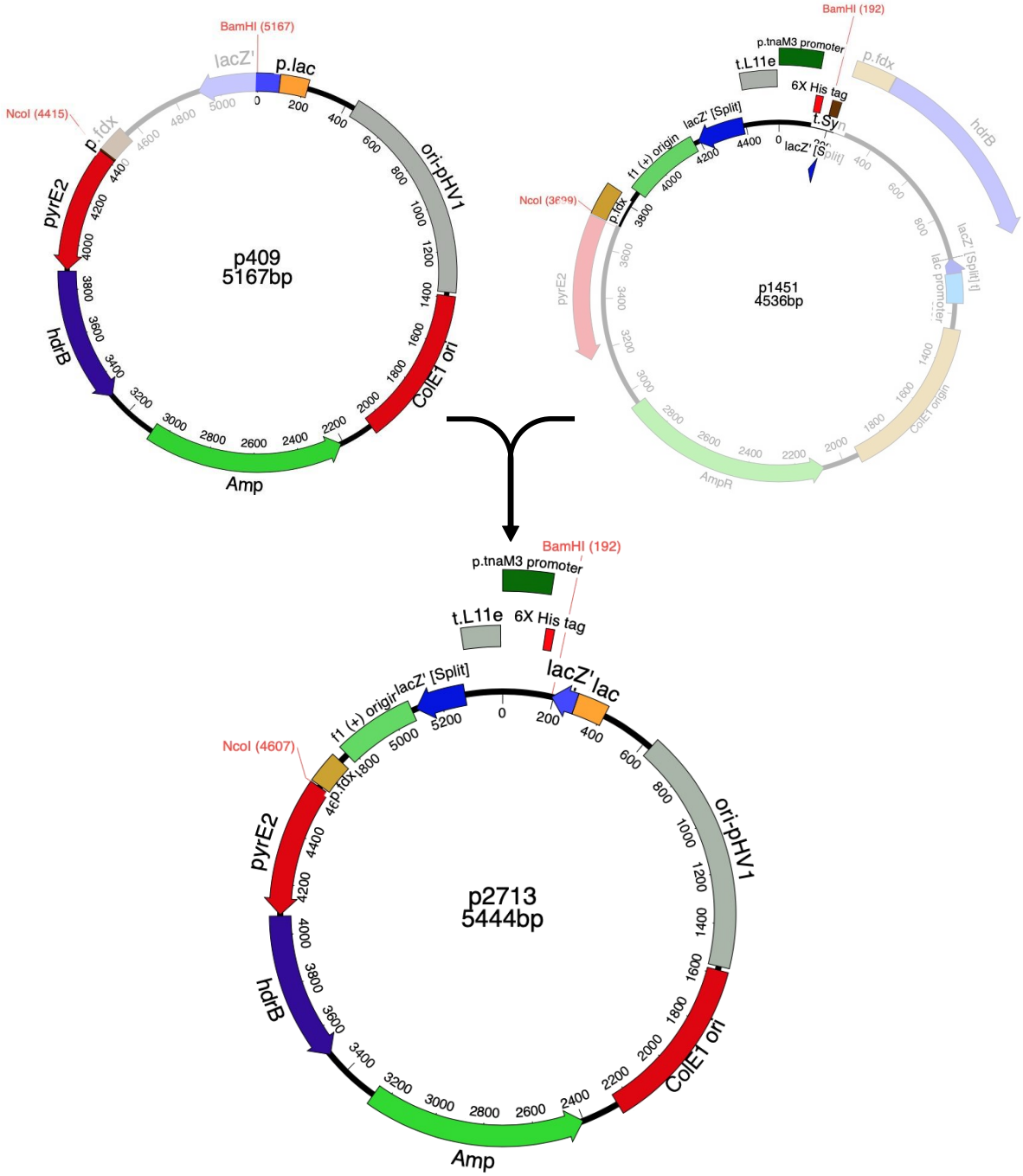


Figure 33. Construction of vector p2713 – the empty vector/'master plasmid'

The "master plasmid" that would supply the backbone and basic features of all the experimental plasmids was produced using the backbone of p409, chosen for its pHV1 origin (which confers low copy number in the cell), as well as its useful marker genes *pyrE2* and *hdrB*. The section added from p1451 between the NcoI and BamHI sites introduced the *p.tna*M3 promoter, which confers lower expression levels than the native *p.tnaA* promoter (Braun et al., 2019). This was chosen alongside the low-copy-number pHV1 origin, as Hel308 is known to be toxic at high levels, and its native level

of expression is thought to be quite low. The transcription terminator *t.L11e* prevents through-transcription from other parts of the plasmid, allowing tighter control of expression (Shimmin & Dennis, 1996). The 6-His tag, used for protein purification, is removed during subsequent cloning steps.

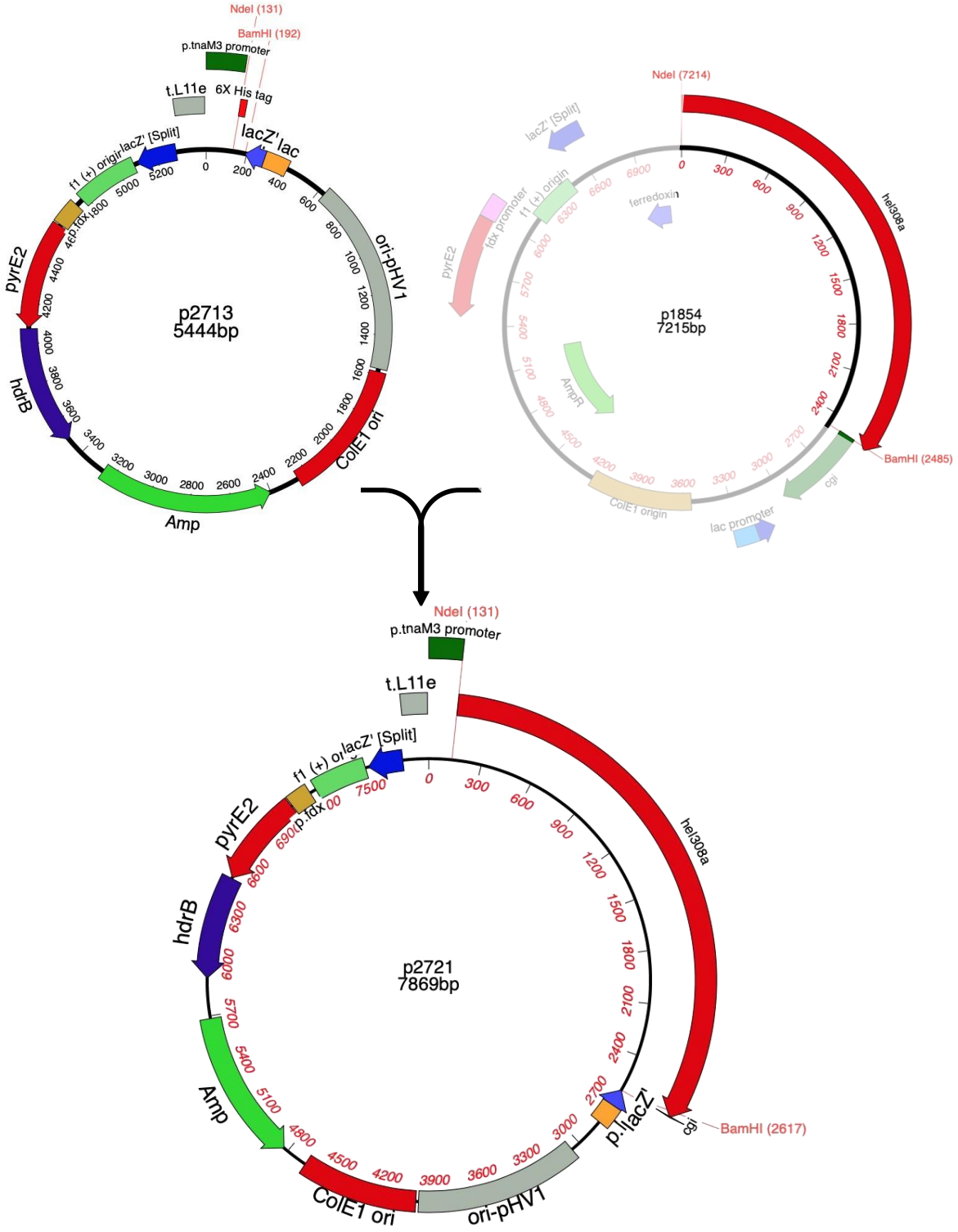


Figure 34. Construction of plasmid for inducible expression of hel308

The *hel308* gene was excised from plasmid p1854, and introduced to p2713 between the NdeI and BamHI sites. This places the start codon of *hel308* directly downstream of the *p.tna*M3 promoter, removing the Histidine tag present in p2713.

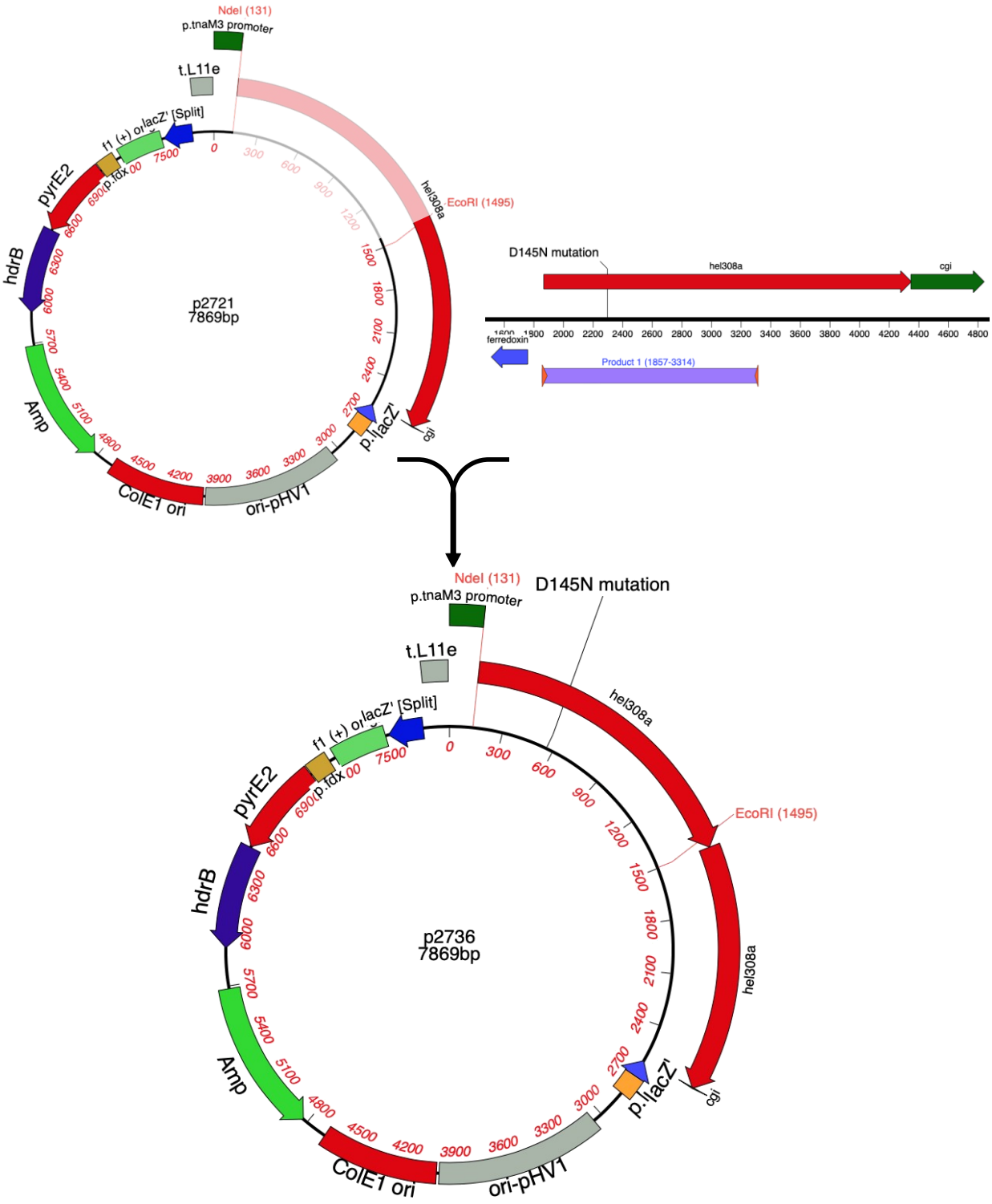


Figure 35. Construction of plasmid for inducible expression of Hel308-D145N

The D145N point mutation was introduced into the p2721 plasmid by cloning part of the *hel308* gene bearing the point mutation from p1335. Using the primers o1645 and o867 introduced the NdeI site at the start codon, allowing fragment swap between the NdeI and EcoRI sites of p2721.

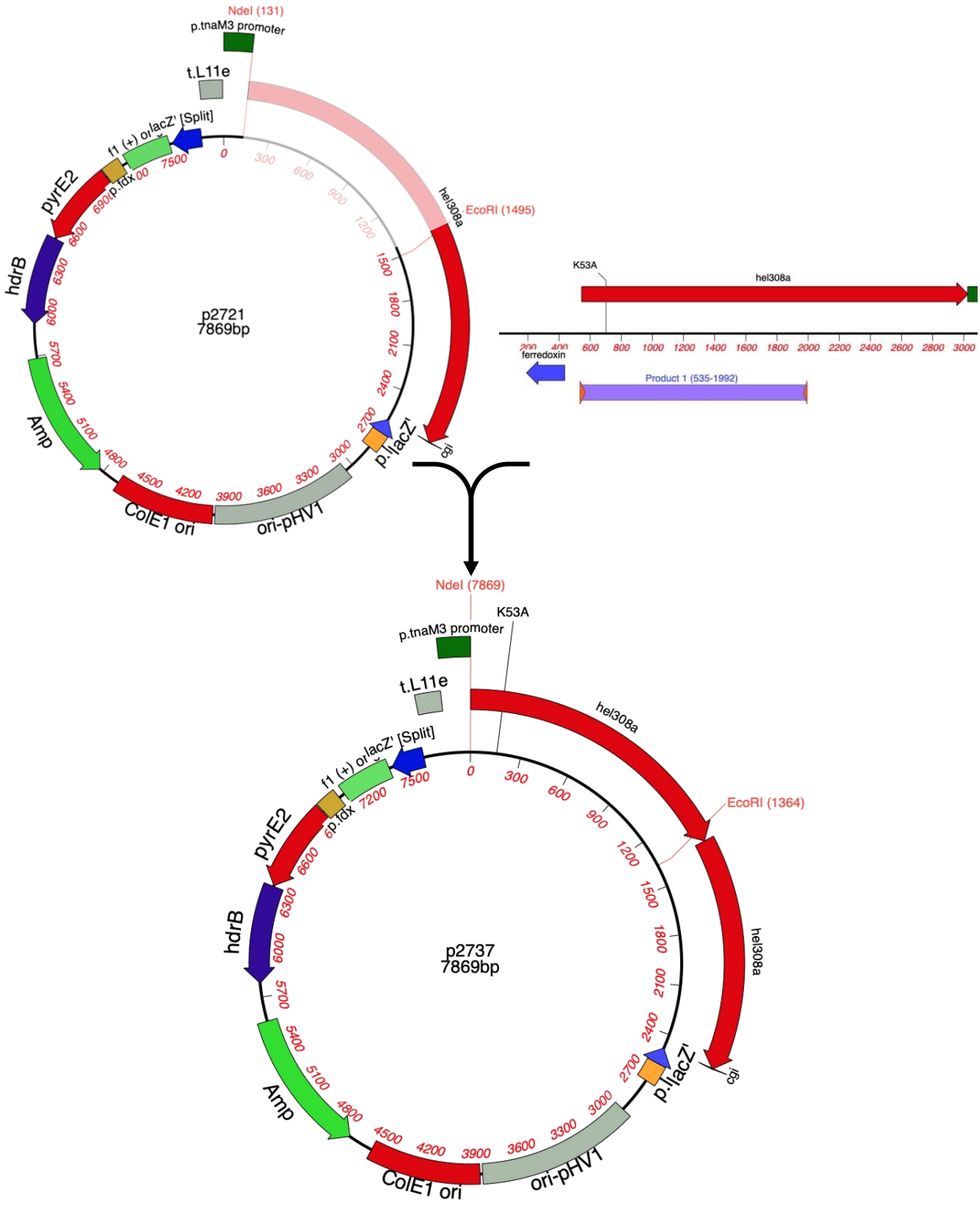


Figure 36 Construction of plasmid for inducible expression of Hel308-K53A

The K53A point mutation was introduced into p2721 by cloning part of the *hel308* gene bearing the point mutation from p1988. Using primers o1645 and o867 introduced the NdeI site at the start codon of the gene, allowing fragment swap between the NdeI and EcoRI sites of p2721. However, the above process was complicated in this case due to the point mutation in question. Previous work (Lever, 2020) had reported difficulty in transforming this gene into *E. coli*, and had hypothesised the lethality of this gene in this species. Placing the gene under the tryptophan-inducible promoter was

unlikely to remedy this, as the promoter does not offer strong repression in *E. coli.* To avoid this *E. coli* lethality, the ligated plasmid DNA was transformed straight into H1208, without the usual passaging in *E. coli* and *dam*$^{-}$ *E. coli* strains. While this would usually not be tolerated well by *H. volcanii* (due to digestion of *dam*-methylated DNA by the Mrr endonuclease), the *mrr* gene has been deleted from the target strain (H1208), allowing good transformation efficiency in this case (Allers et al., 2010).

4.3.2 Strains used in this experiment

Table 4-1 Strains used in this experiment

Strain	Genotype	Notes
H77	*ΔtrpA*	Produced by Thorsten Allers Control strain to confirm tryptophan gradient
H1208	*ΔpyrE2* *Δhdrb* *Δmrr*	Produced by Thorsten Allers *Δmrr* allows transformation with methylated DNA
H5525	*ΔpyrE2* *Δhdrb* *Δmrr* *{pyrE2+ hdrB+}*	Produced by Olivia Wood H1208 + p2713. Empty-vector control
H5524	*ΔpyrE2* *Δhdrb* *Δmrr* *{p.tnaM3::hel308 pyrE2+ hdrB+}*	Produced by Olivia Wood H1208 + p2721. Inducible wild-type *hel308*
H5523	*ΔpyrE2* *Δhdrb* *Δmrr* *{p.tnaM3::hel308-K53A pyrE2+ hdrB+}*	Produced by Olivia Wood H1208 + p2737. Inducible *hel308-K53A*
H5522	*ΔpyrE2* *Δhdrb* *Δmrr* *{p.tnaM3::hel308-D145N pyrE2+ hdrB+}*	Produced by Olivia Wood H1208 + p2736. Inducible *hel308-D145N*

To ensure repression of the potentially lethal gene throughout the transformation process of p2737, the regeneration step was carried out in a Cas-based broth, which lacks tryptophan, rather than the usual YPC regeneration broth. This was to ensure that the gene would not be expressed prior to the experiment.

Verification of gene expression had previously been carried out by transforming p2721 into a *Δhel308* strain (H1391). In the presence of tryptophan, expression of the wild-type Hel308 protein was able to complement the growth defect observed in *Δhel308* strains. This growth assay is shown in a previous chapter; see section 3.4.3. As all of the plasmids used differ only in the presence or absence of the point mutants, it can be assumed that they are equally functional in terms of expression.

4.3.3 Results

Tryptophan gradient plates were produced, containing plain Cas media rather than Cas-uracil. This was due to the noted tendency of *H. volcanii* to spontaneously lose or recombine plasmids when under pressure, as reported in Lever's (2020) attempt to examine K53A point mutants. The lack of uracil in the media ensures that cells not containing the *pyrE2*-marked plasmid will not be able to grow, and so will not interfere with the results.

All of the produced strains which had been transformed with the plasmids listed above were painted across the gradient plates in salt water solution with an autoclaved paintbrush. H77 was also included, as this strain is a tryptophan auxotroph, and will serve to visualise the gradient. The plates were then incubated in the 45°C incubator for four days, allowing colonies to develop.

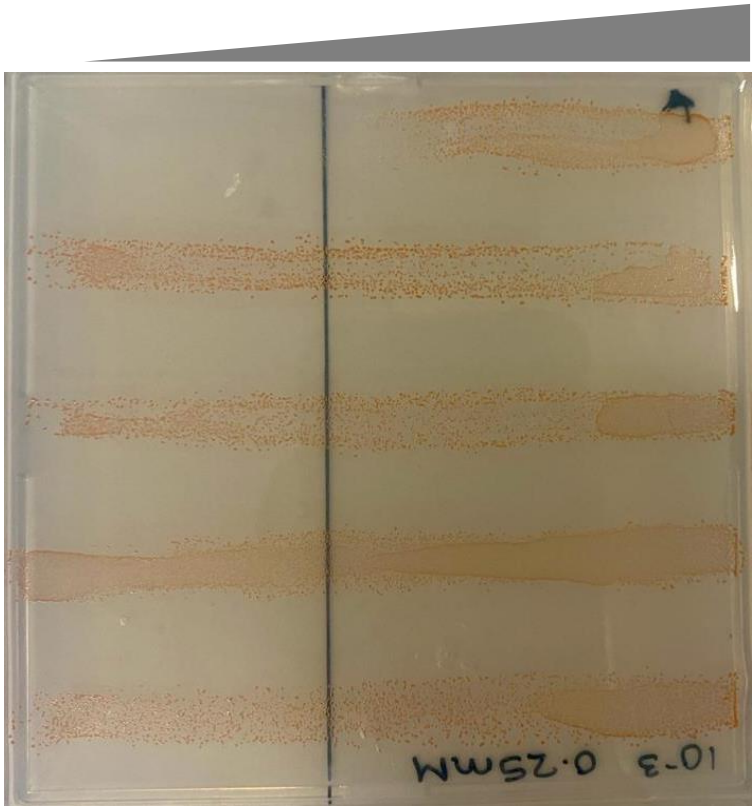


Figure 37. Tryptophan gradient plate showing the growth of the five strains. Concentration of tryptophan increases from left to right, with the black line indicating the narrow edge of the tryptophan wedge. Image courtesy of Olivia Wood.

As can be seen in the above figure, the tryptophan gradient is clearly present on the plate, as visualised by H77. However, the gradient does not appear to have any effect on the growth of the other strains, which grow all the way across the plate, even in the presence of the highest concentration of tryptophan. The presence of the plasmid is indicated by the growth of these strains in the absence of uracil, but as an additional verification, individual colonies were picked and extracted plasmids sent for sequencing to confirm the presence of the point mutations. Both were found to be present, indicating that the Hel308-K53A protein is non-lethal in this context.

4.4 K53A lethality in *Δhel308* strain

4.4.1 Rationale

The experiment that reported K53A's lethality explored the expression of this point mutant in a H4181 background – a strain which is deleted for *hel308, pyrE2* and *hdrB* (Lever, 2020), which could have affected the outcome of the experiment. It had been previously hypothesised that the lethality induced by *hel308-K53A* could be due to affected helicase binding the DNA, but then being unable to translocate. This immobile helicase could prove difficult to remove, thus causing a 'road block' to other enzymes which move along the DNA strand, such as those involved in DNA replication or transcription.

As wild-type Hel308 has been shown to remove bound proteins and complexes from the DNA (Richards et al., 2008), it was considered possible that the presence of wild-type Hel308 in the above experiment could be "rescuing" the cell, and thus masking the damaging effects of this point mutation. This could explain the discrepancy observed between Lever's original results, obtained from a *Δhel308* strain, and the initial results in the H1208 background. The same assay was therefore repeated in a *Δhel308* strain.

4.4.2 Strain construction

As a follow-up to the experiments above, the *hel308* gene was deleted from H1208 using the deletion vector p1254, and the deletion confirmed by Southern blot, shown in Figure 38, below.

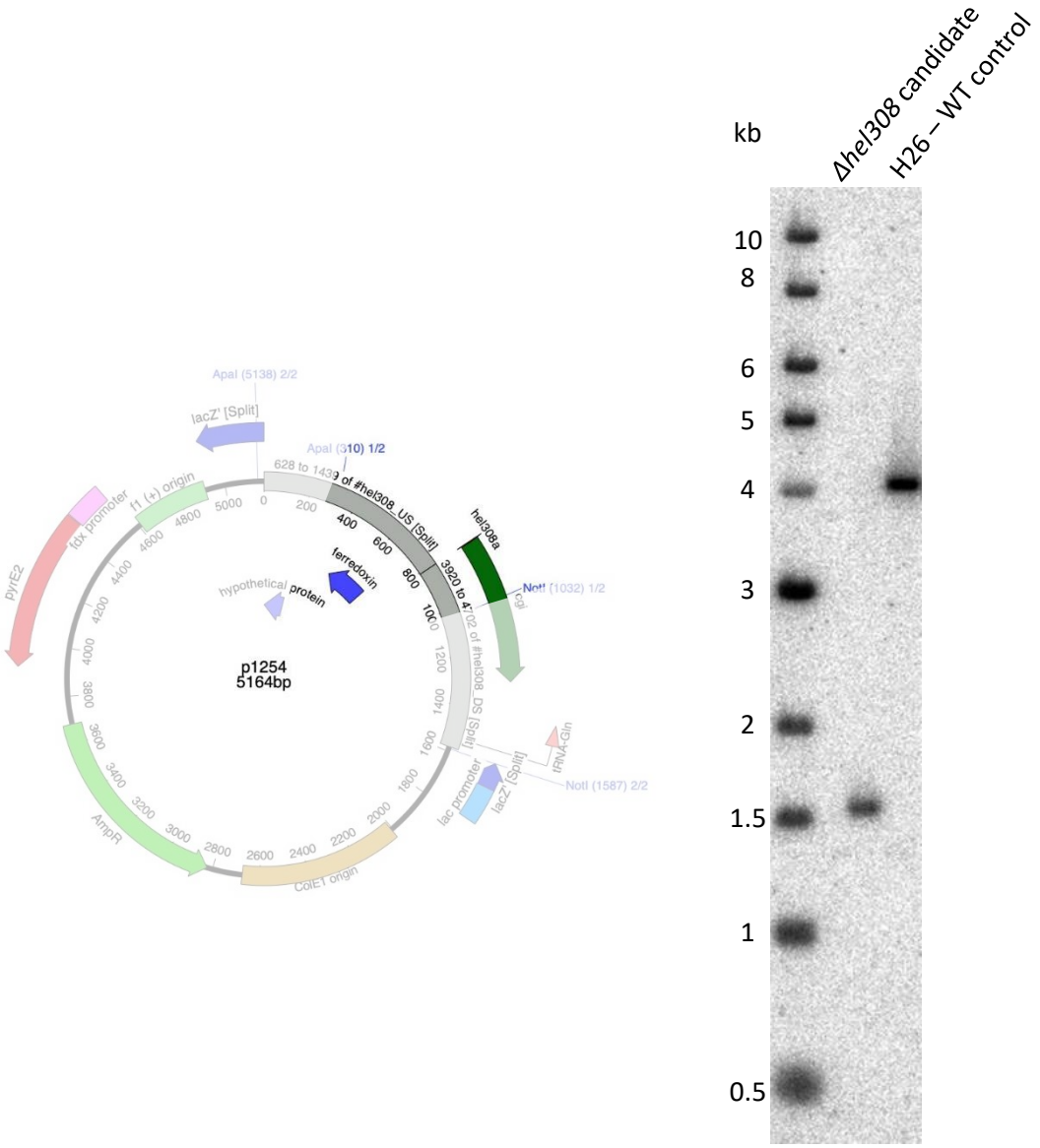


Figure 38 Southern blot confirming the deletion of hel308 *from H5452. Genomic DNA had been digested with NcoI and MluI. Probe used is shown. Expected band sizes were 4kb for wild-type* hel308*, 1.5kb for* Δhel308*.*

The new strain, bearing the genotype *ΔpyrE2, ΔhdrB, Δmrr, Δhel308*, was named H5453. The same plasmids were then transformed into H5453. List of strain numbers and their genotypes are shown below.

Table 4-2 Strains used in this section

Strain	Genotype	Notes
H5452	*ΔpyrE2* *ΔhdrB* *Δmrr* *Hel308+*::[*Δhel308 pyrE2+]*	Produced for this experiment. Pop-in of p1254, in preparation to delete chromosomal *hel308*
H5453	*ΔpyrE2* *ΔhdrB* *Δmrr* *Δhel308*	Produced for this experiment. Chromosomal *hel308* deleted by pop-out; confirmed by Southern blot.
H5542	*ΔpyrE2* *ΔhdrB* *Δmrr* *Δhel308* *{pyrE2+ hdrB+}*	Produced for this experiment. H5453 + p2713. Empty vector control
H5543	*ΔpyrE2* *ΔhdrB* *Δmrr* *Δhel308* *{p.tna*M3::*hel308 pyrE2+ hdrB+}*	Produced for this experiment. H5453 + p2721. Inducible wild-type *hel308*
H5544	*ΔpyrE2* *ΔhdrB* *Δmrr* *Δhel308* *{p.tna*M3::*hel308-D145N pyrE2+ hdrB+}*	Produced for this experiment. H5453 + p2736. Inducible *hel308-D145N*
H5545	*ΔpyrE2* *ΔhdrB* *Δmrr* *Δhel308* *{p.tna*M3::*hel308-K53A pyrE2+ hdrB+}*	Produced for this experiment. H5453 +p2737. Inducible *hel308-K53A*

4.4.3 Results

The constructed and verified strains were then grown on tryptophan gradient plates, as described in the Methods chapter (2.2.7). Plates were examined following five days of growth.

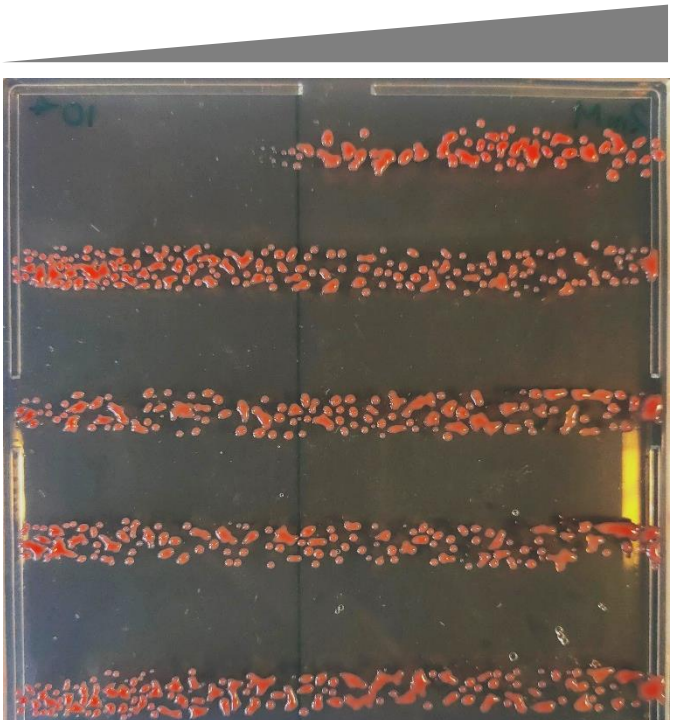


Figure 39 Tryptophan gradient plate showing the growth of the five strains. Concentration of tryptophan increases from left to right to a maximum of 2mM, with the black line indicating the narrow edge of the tryptophan wedge.

Following five days of growth, it was clear that none of the *hel308* point mutants caused lethality when induced by tryptophan. Colonies from the high-tryptophan side of the plate were subjected to sequencing to confirm the presence of both the promoter and the point mutation. Both were found to be present and intact. This confirms the non-lethality of Hel308-K53A in strains both with and without WT-Hel308.

4.5 Chromosomal replacement of *hel308-K53A*

4.5.1 Strain construction

Following the revelation that *hel308-K53A* is not a lethal gene, a decision was made to attempt to replace the chromosomal copy of *hel308* with *hel308-K53A*. This would allow better characterisation of the effects of the allele, opposed to its inducible expression through an episome. Introduction of the mutant allele into the *hel308* chromosomal locus would allow it to be expressed and controlled by its native promoters. A plasmid to perform this replacement already existed – p1988, produced by B. Lever (*dam*- preparation p2015). This plasmid is shown below.

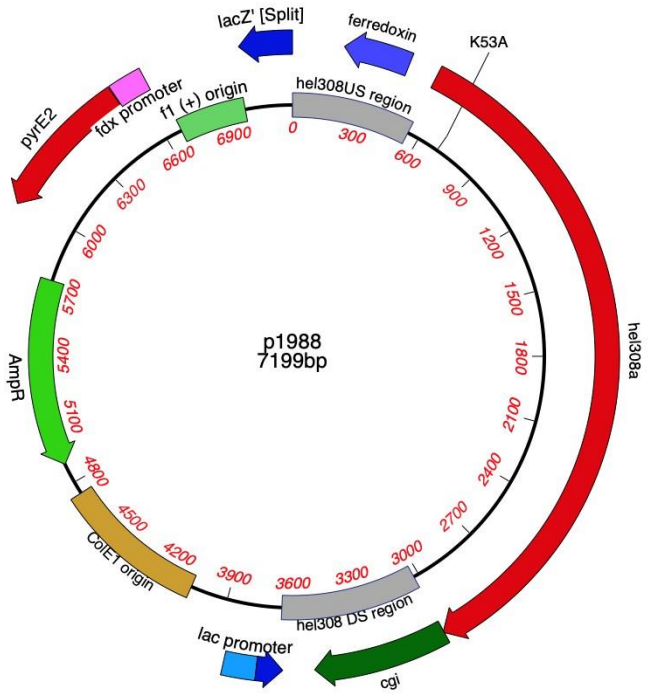


Figure 40. p1988, produced by B. Lever. Vector for chromosomal replacement of hel308 *with* hel308-K53A.

The plasmid was transformed into the host strain H1393 (*ΔpyrE2, ΔtrpA, Δhel308::trpA+*) to produce pop-in H5634. Pop-outs were generated as per the Methods chapter, and were plated onto Cas+5FOA+trp. Once colonies had grown, individual colonies were restreaked onto both Cas+ura and Cas+ura+trp. Colonies that grew on Cas+ura+trp, but not on Cas+ura, were thought to have lost the *trpA* marker from all copies of the chromosome, thus reverting them to tryptophan auxotrophy. Five candidates with these qualities were prepared for further evaluation by Southern blot. Results of the Southern blot are shown below.

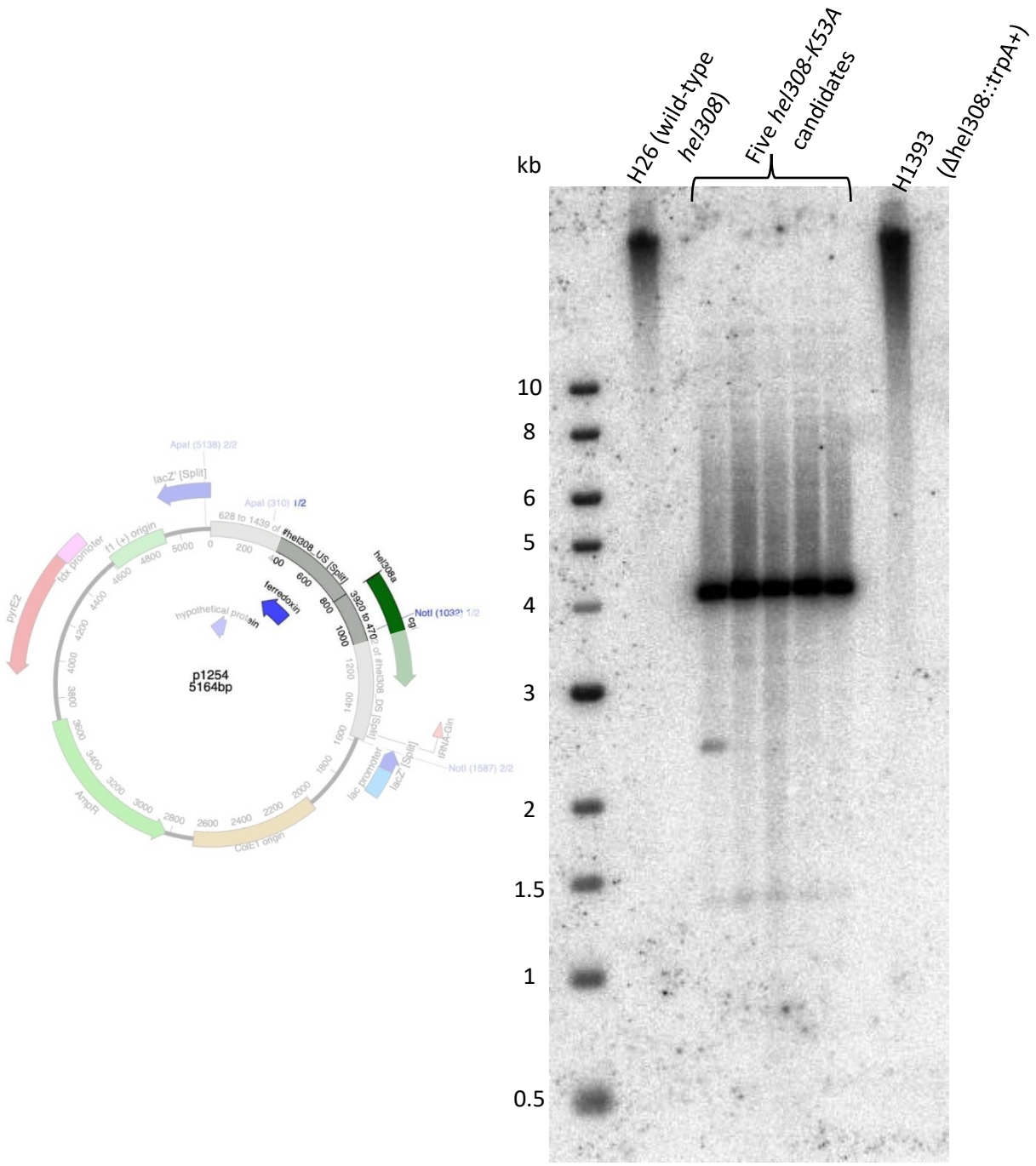


Figure 41. Probe construction (left) and Southern blot confirming the presence of hel308-K53A *in the selected candidates (right). Genomic DNA was digested with NcoI and MluI. Note that candidate 1 shows a faint band at 2.4kb, corresponding to the* Δhel308::trpA+ *allele.*

Unfortunately, both control lanes misran during production of the Southern blot, appearing as dark smudges at the top of the membrane. This is likely due to poor resuspension of the genomic DNA following precipitation. However, of the candidates trialled here, all five bear the *hel308-K53A* allele (the 4kb band) which cannot be differentiated from wild-type using this method. However, candidate 1 also shows a faint band at 2.4kb, corresponding to the *Δhel308::trpA+* allele. This candidate was therefore not used for further experiments, owing to its merodiploid nature. While this Southern is not ideal due to the failure of the controls to run, the presence of the desired band at the correct height, coupled with the phenotype of trp-, was considered sufficient evidence to

accept candidate 2 and progress with characterisation of the point mutant. With the point mutant confirmed by sequencing, the new strain was named H5721 (*ΔpyrE2, ΔtrpA, hel308-K53A)*.

4.5.2 Strain characterisation

A growth assay was carried out to assess the relative growth rate of H5721 in YPC compared to strains with wild-type *hel308* (H53), *Δhel308* (H2085), and *hel308-D145N* (H1555). Aside from variation of the *hel308* allele, all strains possess the same genetic background (*ΔpyrE2, ΔtrpA*). Assay was carried out as described in the Methods chapter. Ten biological replicates were included per strain.

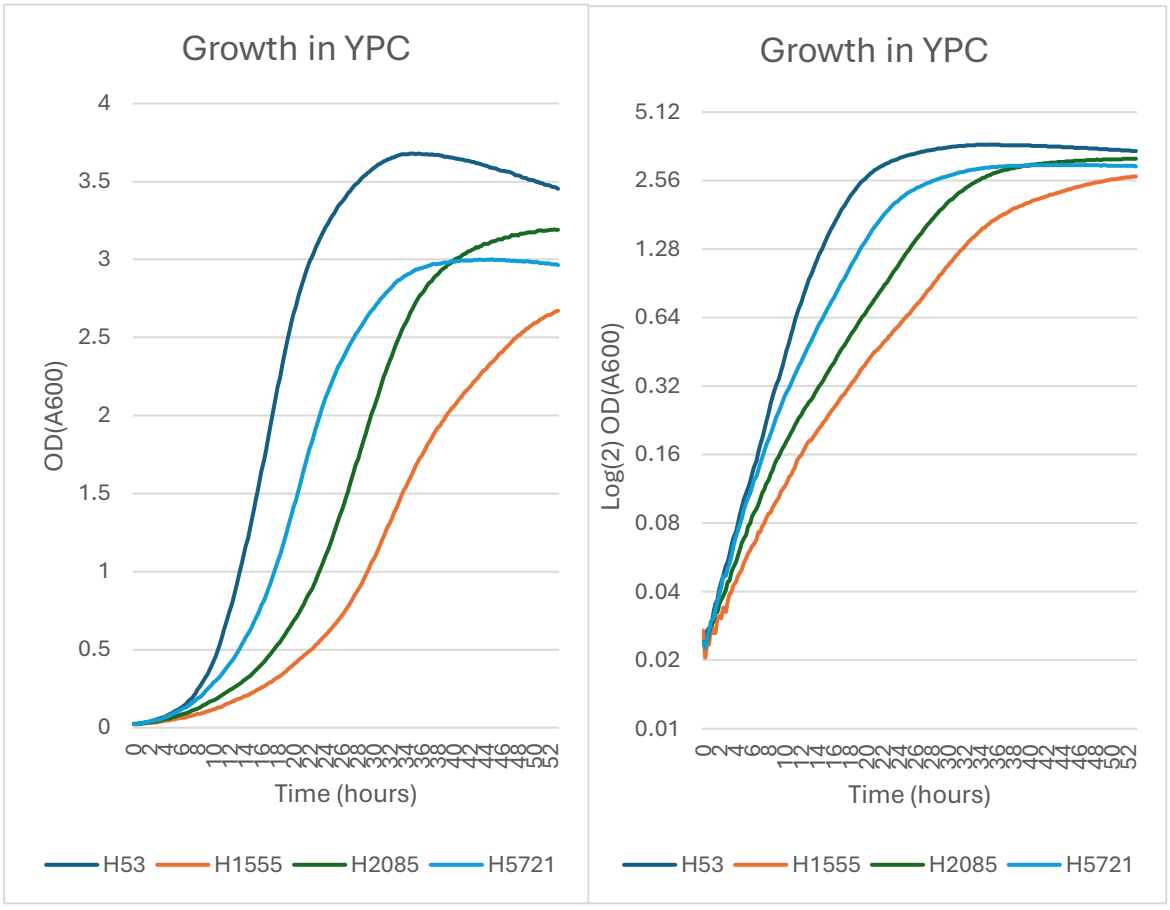


Figure 42 Growth rate of the four strains in YPC. Data are shown using a linear y-axis (left) and a logarithmic y-axis (right)

As can be seen in Figure 42, H53 maintains a clear growth advantage over the other strains. However, it is notable that differences are clearly observable between the three remaining strains. Not only are the two ATPase-null point mutants clearly different from the *Δhel308*, but they are also different from each other.

In addition to the marked differences between growth rate and stationary OD observed between the four strains, in the graph with the logarithmic y-axis, some biphasic growth may be evident. The growth rate of the strains seems to be steeper earlier, and then slows slightly into a second phase before transitioning into stationary phase. Doubling times were calculated for both of these exponential phases for all four strains, as shown in the table below.

Table 4-3 Calculated doubling times of the strains used above

	Early exponential	Slower exponential
H53	2.5h	2.75h
H1555	5.25h	7.25h
H2085	4.5h	5.5h
H5721	3h	4.5h

While the doubling times themselves vary, the relationship between the four seems roughly consistent in both phases. H1555 (*hel308-D145N*) seems to be consistently the slowest growing of the strains investigated, which seems unusual given that the two ATPase-null point mutants were previously expected to produce similar phenotypes. All three mutant alleles trialled conferred a growth defect compared to the wild type. Interestingly, it appears that the phenotypes conferred by the ATPase-null point mutants are not equivalent that produced by the *Δhel308* phenotype. This suggests that the mutant Hel308 is still having some effect within the cell, and may be involved in some activity or interaction which does not require ATPase activity.

In addition, the phenotypes produced by the two ATPase-null point mutants do not appear to be equivalent to each other, which is unexpected. However, the modes by which they are expected to be ATPase-null differ. Hel308-D145N is predicted to be unable to co-ordinate a Mg^{2+} ion necessary for ATP binding, while Hel308-K53A is predicted to be able to bind ATP but not hydrolyse it. It could be that the binding of ATP alters the conformation of the protein; affecting its interactions with other molecules in the cell. There could potentially be a scenario (suggested in Lever, 2020) in which ATPase-null Hel308 is able to bind DNA, but not translocate; thus forming a potential road-block to other proteins or complexes which translocate along DNA strands.

Hel308 has previously been implicated in DNA repair (Richards et al., 2008), so a UV irradiation assay was carried out to determine the effects of the *hel308-K53A* allele compared to wild-type, *Δhel308,* and *hel308-D145N*. UV assay was carried out as described in the Methods chapter, and photoreactivation avoided by storing the plates in a black bag immediately following irradiation. The data are shown below.

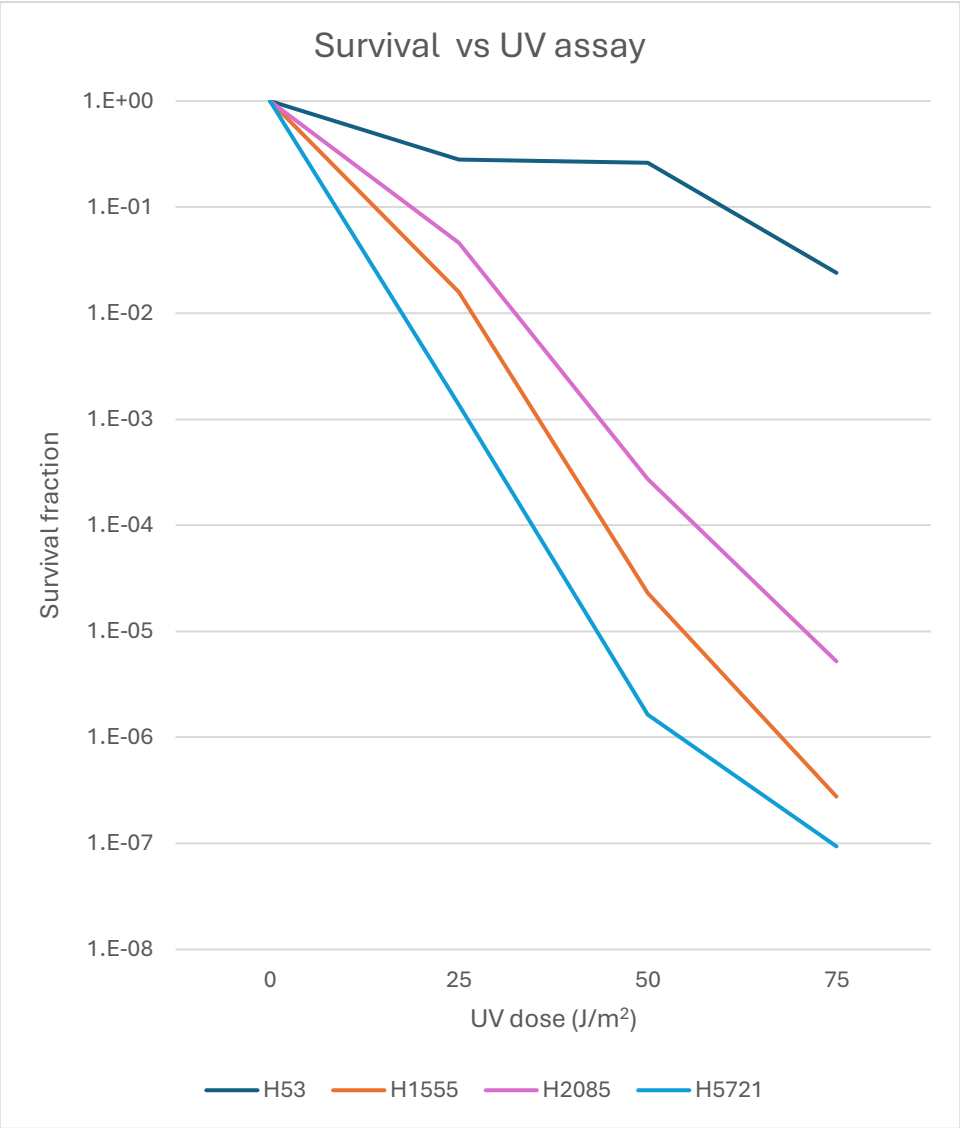


Figure 43. Survival frequency of strains bearing different hel308 *alleles following exposure to UV damage. Survival is expressed as a fraction relative to growth without UV exposure. Each data point is the result of three biological replicates.*

As can be seen in Figure 43, H53 shows a much higher resistance to UV than the other three strains, suggesting that *hel308,* and specifically its helicase ability, is highly important in the cell response to this variety of DNA damage. In this regard, the two ATPase-null point mutants do not appear to exhibit markedly different phenotypes.

It has been previously reported that the *hel308-D145N* point mutant produces a different phenotype to the *Δhel308* phenotype (Lever et al., 2023). Both *Hel308-D145N* and *Δhel308* cause sensitivity to mitomycin C; however, *hel308-D145N* causes a small decrease in recombination, while *Δhel308* increases recombination 5.5-fold. Unfortunately, the effects of *hel308-K53A* on mitomycin C sensitivity could not be explored due to a shortage of this chemical. It is however one of the strains tested against aphidicolin in a later chapter (see Chapter 6.4).

4.6 Future exploration

Full characterisation of the Hel308-K53A protein could not be performed due to time constraints. However, there are many experiments that could be performed now that this allele is confirmed as non-lethal.

4.6.1 Effects on recombination rate

In the current study, the strain bearing the *hel308-K53A* chromosomal replacement was not tested for recombination rate, as the point mutant had not been transformed into an appropriate background for this assay. Its effects on recombination rate within the cell therefore cannot be verified. While it may seem likely that the effects on recombination rate will be similar to that of the other ATPase-null point mutant *hel308-D145N*, it may be worthwhile to verify this. It is possible that Hel308-K53A, which is predicted to be unable to bind ATP, may behave differently to Hel308-D145N, which can bind ATP but cannot hydrolyse it. In fact, the differences in exponential growth rate between the two strains (H5721 and H1555) suggest that the two proteins may have slightly different effects or interactions.

The typical method for determination of recombination rate would be replacement of this allele on the chromosome of H164 (*ΔpyrE2, ΔtrpA, bgaHa-Bb, leuB-Ag1).* This background bears a *leuB* allele containing a frameshift mutation near the end which renders it non-functional. This strain is thus leucine auxotrophic. In the assay, an originless plasmid (p163) bearing a second leucine allele and the *pyrE2* marker is introduced. The leucine allele on the plasmid bears a different frameshift mutation near the other end of the gene. While both alleles are non-functional, recombination between the two results in one complete *leuB* gene and one bearing both mutations. This is detected by leucine autotrophy (Delmas et al., 2009). This assay can also detect the difference between crossover and non-crossover events; where a crossover event has occurred, cells will exhibit the phenotype ura+ leu+, as a result of the whole plasmid recombining onto the chromosome. Strains which have gained a functional *leuB* allele as a result of a non-crossover event (or gene conversion) will not bear a *pyrE2* marker on the chromosome. As the plasmid used in the assay does not contain an origin of its own, it will not be replicated, and will be lost as the culture grows. A schematic of this technique is briefly summarised in Figure 44, below.

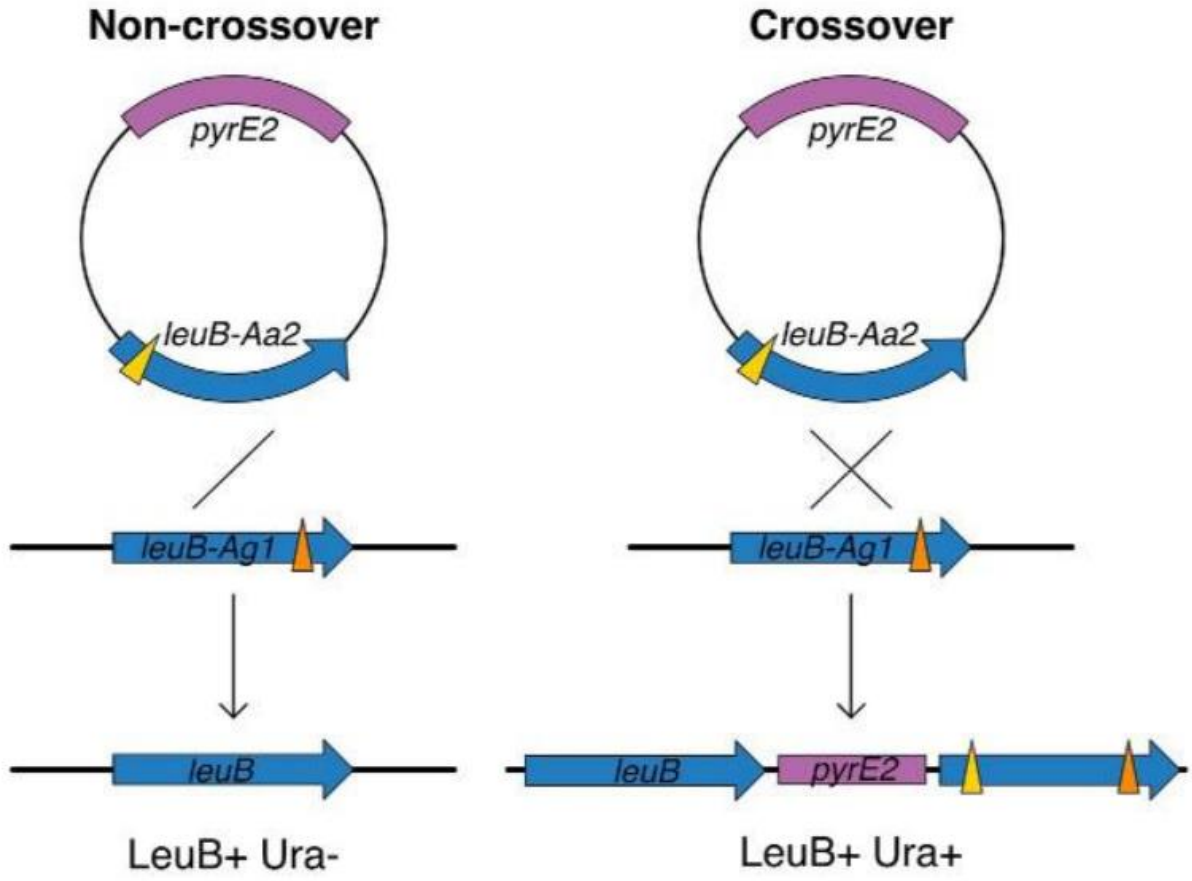


Figure 44. Diagram of the recombination assay. Recombination events between the two mutant leuB *alleles can generate a functional* leuB *allele. Crossover (right) versus non-crossover events (left) can be differentiated by the presence of the* pyrE2 *marker gene on the chromosome. From Lever et al. (2023)*

This aspect of the phenotype of strains bearing the *hel308-K53A* allele is worth exploring, as the other ATPase-null point mutant *hel308-D145N* was found to produce a different phenotype than *Δhel308* (Lever, 2020). The presence of *hel308-D145N* was found to reduce the recombination rate by half compared to the wild type, while deletion of *hel308* increased recombination five-fold. This could suggest an interaction between Hel308 and another protein, which does not require ATPase or helicase activity to take effect. Alternatively, the ATPase-null Hel308 could be binding to its usual substrates but not performing its usual function. It has previously been shown that Hel308 does not require ATP in order to bind DNA forks and begin melting of the duplex (Büttner et al., 2007). It is therefore likely that both ATPase-null mutants are able to bind to DNA but not to translocate along it, as was the case in equivalent point mutants from other species (Fujikane et al., 2006; Guy & Bolt, 2005; Richards et al., 2008). This is unlikely to contribute to DNA repair, and may function as a barrier to alternative pathways.

4.7 Conclusion

It was previously claimed that the ATPase-null *hel308-K53A* point mutation was a lethal gene in both *H. volcanii* and *E. coli*. Expression of this point mutant through an inducible promoter did not cause death in *H. volcanii* however. Further exploration of the same inducible allele in a background similar to that of the originally reported effect also did not cause lethality.

The point mutant was able to be introduced onto the chromosome to replace the *hel308* locus. This mutant strain was briefly characterised, and was found to behave similarly, but not perfectly equivalent to, other ATPase-null point mutant *hel308-D145N*.

5. Essentiality of *hel308* in *H. mediterranei*

5.1 Background

5.1.1 H. volcanii and H. mediterranei

H. volcanii and *H. mediterranei* are two closely related haloarchaea that have found success as model organisms. These two species share 86.6% sequence identity between the standard lab strains (Naor et al., 2012). In fact, the *H. volcanii* genome sequence was used in assembling the *H. mediterranei* genome sequence (Han et al., 2012). Maximum likelihood phylogenetic trees showing the evolutionary relationship between these two organisms, and their relationship to other members of the genus, are shown in Figure 45, below.

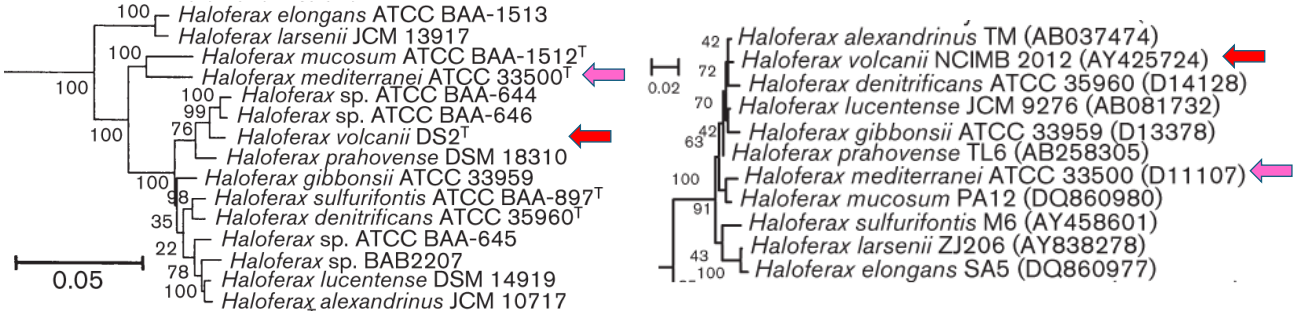


Figure 45. Phylogenetic trees of the Haloferaceae*, based on conserved protein sequences (left, bar: 0.05 substitutions per nucleotide position) and 16S rRNA sequences (right, bar: 0.02 substitutions per nucleotide position). Positions of* H. volcanii *are indicated by the red arrow;* H. mediterranei *is indicated by the pink arrow. From Gupta et al. (2015).*

While the main chromosomes of the two organisms are highly similar, the additional replicons are highly divergent. The one plasmid and three minichromosomes found in *H. volcanii* do not seem to share homology with the three minichromosomes of *H. mediterranei* (Han et al., 2012). In *H. mediterranei,* these additional replicons are denoted HM100 (130 kb) HM300 (322 kb) and HM500 (504 kb). While *H. volcanii* has been used extensively as a model organism to examine mechanisms of archaeal DNA repair and replication (reviewed in Pérez-Arnaiz et al. (2020)), *H. mediterranei* has garnered more attention for its potential use in biotechnology. Many species of haloarchaea produce polyhydroxyalkanoates (PHA) as an intracellular carbon source (Koller, 2019). These are polymers of 3-hydroxbutyrate that form water-insoluble granules within the cell, and can be harvested for production of biodegradable plastics as an alternative to fossil carbon sources (reviewed in Mitra et al. (2020)). While small amounts of PHA have been extracted from *H. volcanii* cultures (Fernandez-Castillo et al., 1986), effective production in this organism requires the addition of enzymes from other species (Han et al., 2009). Meanwhile, *H. mediterranei* has proved highly effective in PHA production, with PHA content of over 70% of cell weight possible, depending on the growth conditions and carbon sources utilised (Mitra et al., 2020). Much of the research into *H. mediterranei* has therefore focused on this aspect of its metabolism, rather than other details of its biology.

In addition to differences in PHA production, the DNA replication mechanisms in the two species also appear to vary. All of *H. volcanii*'s chromosomal origins can be deleted without impairing growth rate; cells instead exhibit an increased growth rate in the absence of these origins (Hawkins et al., 2013). Meanwhile, deletion of *H. mediterranei*'s chromosomal origins results in activation of a dormant origin (Yang et al., 2015). Deletion of this origin is not possible where the other origins have

already been removed; it seems that deletion of any three of the four is tolerated, but that this species cannot survive in the absence of origins. The mechanisms influencing essentiality of origins in these two species have not yet been identified, but may be affected by differences in frequency of homologous recombination between the two organisms (Dattani et al., 2022).

5.1.2 Hel308 homologues

It was previously claimed that *hel308* is essential in *H. mediterranei*, which is not the case in *H. volcanii* (Dattani, unpublished research). It was considered unusual that this should be the case, when the two organisms, and indeed the two genes, are so similar. Hel308 is not thought to be essential in most of the archaeal species that encode it, although it has been found to be essential in the *Sulfolobus* species *tokodaii* and *islandicus* (Hong et al., 2012; Zhang et al., 2013). However, it could suggest that Hel308 plays a role in origin-independent replication, contributing to the contrasting behaviour of these two organisms.

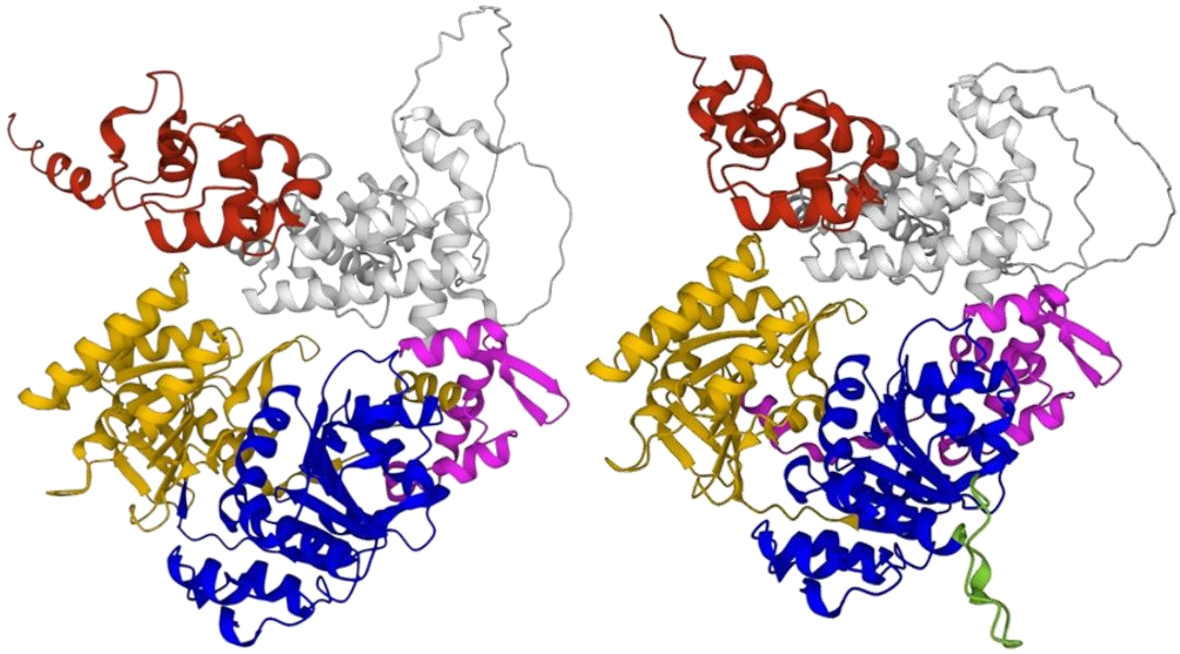


Figure 46. Structure of H. volcanii *Hel308 (left) and* H. mediterranei *Hel308 (right). Domains are coloured as follows: domain 1 (RecA; blue), domain 2 (RecA; orange), domain 3 (winged helix; pink), domain 4 (ratchet; white) and domain 5 (HLH; red). The 18 amino-acid proline-rich loop in domain 1 is present only in the* H. mediterranei *homologue, highlighted in green. Structures courtesy of Alphafold.*

Predicted structures of the two homologues were produced through Alphafold. As can be seen in Figure 46, the two proteins' predicted structures are highly similar, and share 88% amino acid identity and 92% positive identity (amino acids with similar properties). While there are several differences between the two in the ratchet domain, the largest and most obvious difference between the two proteins is located in domain 1. *H. mediterranei* Hel308 contains an inserted stretch of 18 amino acids not seen in *H. volcanii* (G68-T85, shown in green, above, and in alignment, in Figure 47, below). This seems to be a rare structure among *Haloferax* species, being absent from the species *gibsonii, denitrificans, marisrubi, mucosum, alexandrinus* and *sulfurifontis*, but is present in *Haloferax larsenii*. Unfortunately, no research has yet been published on the essentiality of either

hel308 or replication origins in *H. larsenii*, so the significance of this insert is not immediately apparent.

```
Haloferax mediterranei  1    MRTADLTGLPTGIPEALHDEGIEELYPPQAEAVEAGLTDGESLVAAVPTASGKTLVAELA  60
Haloferax volcanii      1    MRTADLTGLPTRIPEALRDEGIEELYPPQAEAVEAGLTDGESLVAAVPTASGKTLVAELA  60

Haloferax mediterranei  61   MLSSVARGVPEPRSDGSGELPSGSTGGKALYIVPLRALASEKKAEFERWEEYGIDVGVST  120
Haloferax volcanii      61   MLSSVARG-----------------GKALYIVPLRALASEKKAEFERWEEYGIDVGVST  102
```

Figure 47. Alignment of the first 100/120 residues of the Hel308 sequence of the two Haloferax *species. Species' names are listed on the left.*

Predictors of protein structure consistently place this structure on the external surface of the protein, possibly in a position to form an interaction with another protein. The 18 amino acids in this insertion are intrinsically disordered, and also include three prolines. Regions high in proline are often the sites of protein-protein interactions (Ball et al., 2005); likewise, the flexibility of intrinsically disordered regions can lend itself to a large number of interaction partners (Babu, 2016). As an explanation for the markedly different properties reported between two very similar proteins, this is a tempting proposition. While the Alphafold structure shown in Figure 46 shows a helical structure for this region, this was predicted with very low confidence (pLDDT < 50), and so may not be an accurate reflection of its structure *in-vivo*.

To explore the essentiality of *H. mediterranei hel308*, an experiment was proposed to explore which qualities of this protein are essential in *H. mediterranei*. This would be carried out by attempting to delete *hel308* from *H. mediterranei* while complementing the deletion with various *hel308* homologues or variants. For example, complementing with *H. mediterranei hel308* on a plasmid should maintain expression levels, allowing deletion of the chromosomal copy of the gene.

In addition, other alleles would be investigated to see whether they were sufficient for complementation in this scenario. Initial alleles selected were *H. volcanii hel308*, and *H. mediterranei hel308*. These were selected as a starting point, with the intention to progress to examining point mutants, or “hybrid genes” composed of different domains of the two homologues spliced together. Once the qualities determining *hel308* essentiality in *H. mediterranei* have been determined, further study aimed to replace *H. mediterranei hel308* with its *H. volcanii* homologue, and the ascertain whether this would allow origin-independent replication in this species.

5.2 Aims and Objectives

- Determine what qualities of *hel308* are essential in *H. mediterranei*
- Attempt to delete *hel308* from multi-origin-deleted *H. mediterranei* to ascertain whether this allows deletion of the final origin.

5.3 *Hel308* complementation assay

5.3.1 Plasmid construction

A plasmid suitable for complementation with *H. mediterranei hel308* already existed – p2680, which had been constructed by A. Dattani, shown in Figure 48, below.

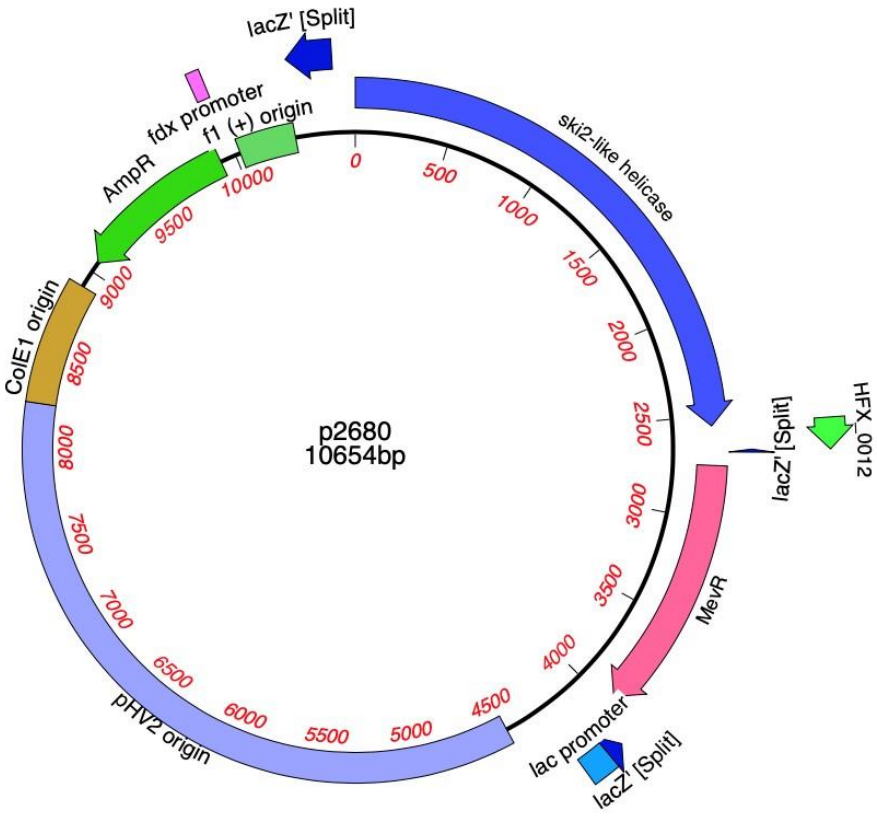


Figure 48. p2680 complementation plasmid; H. mediterranei hel308 *gene is labelled as ski2-like helicase.*

This plasmid contains the pHV2 origin, which is functional in *H. mediterranei* (T. Allers, unpublished research). The *pyrE2* marker could not be used, as this was to be used to mark the chromosomal *hel308* deletion vector, so the selectable marker for this plasmid is instead the mevinolin resistance gene denoted *mevR.* This gene encodes the enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase under a strong promoter, cloned from *Haloarcula hispanica* to avoid recombination with the native homologue on the chromosome. Mevinolin competitively inhibits this enzyme, which blocks production of isoprenoid lipids, which are a major component of the cell membrane (Wendoloski et al., 2001). In wild-type cells, this causes cell death via membrane disintegration. The increased production of HMG-CoA reductase conferred by the *mevR* marker is sufficient to provide resistance to mevinolin.

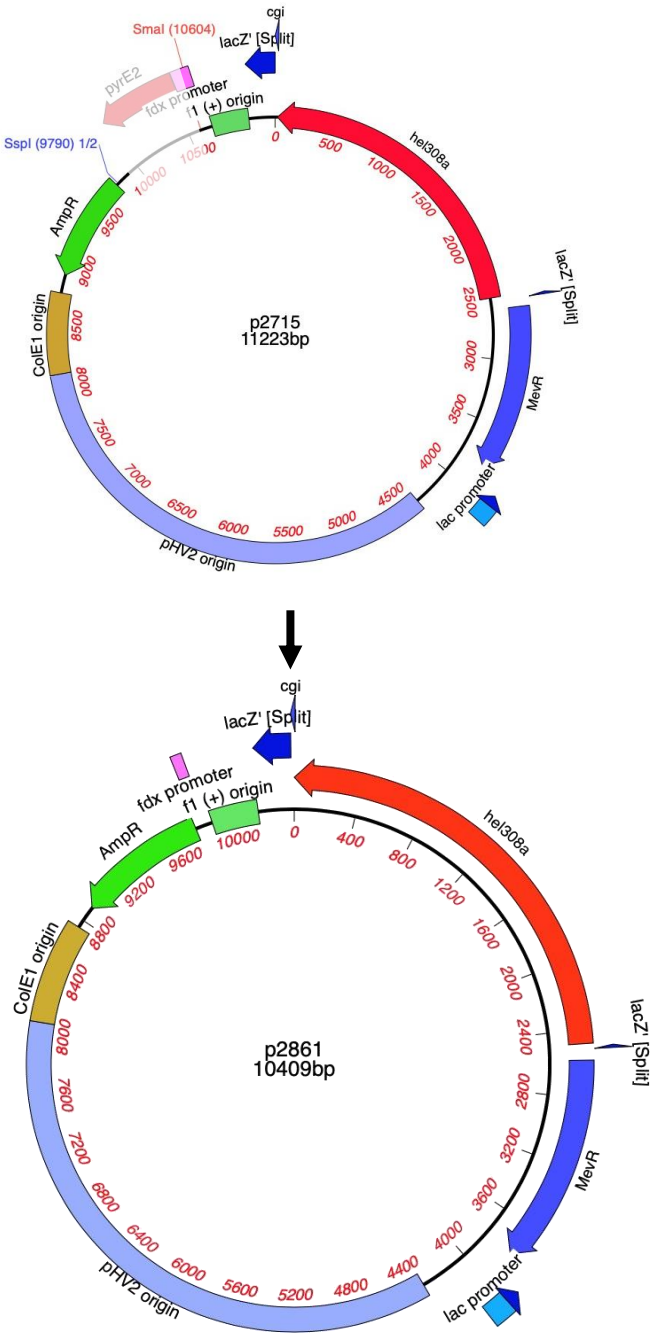


Figure 49. Construction of H. volcanii hel308 *complementation plasmid.*

To produce a *H. volcanii hel308* complementation vector equivalent to p2680, the *pyrE2* marker was removed from p2715 by digestion at the SmaI and SspI sites. The plasmid was then re-ligated after blunt-ended digestion. It should be noted that this results in the *H. volcanii hel308* gene being present in the opposite orientation in the plasmid compared to p2860, but this should not affect expression levels greatly.

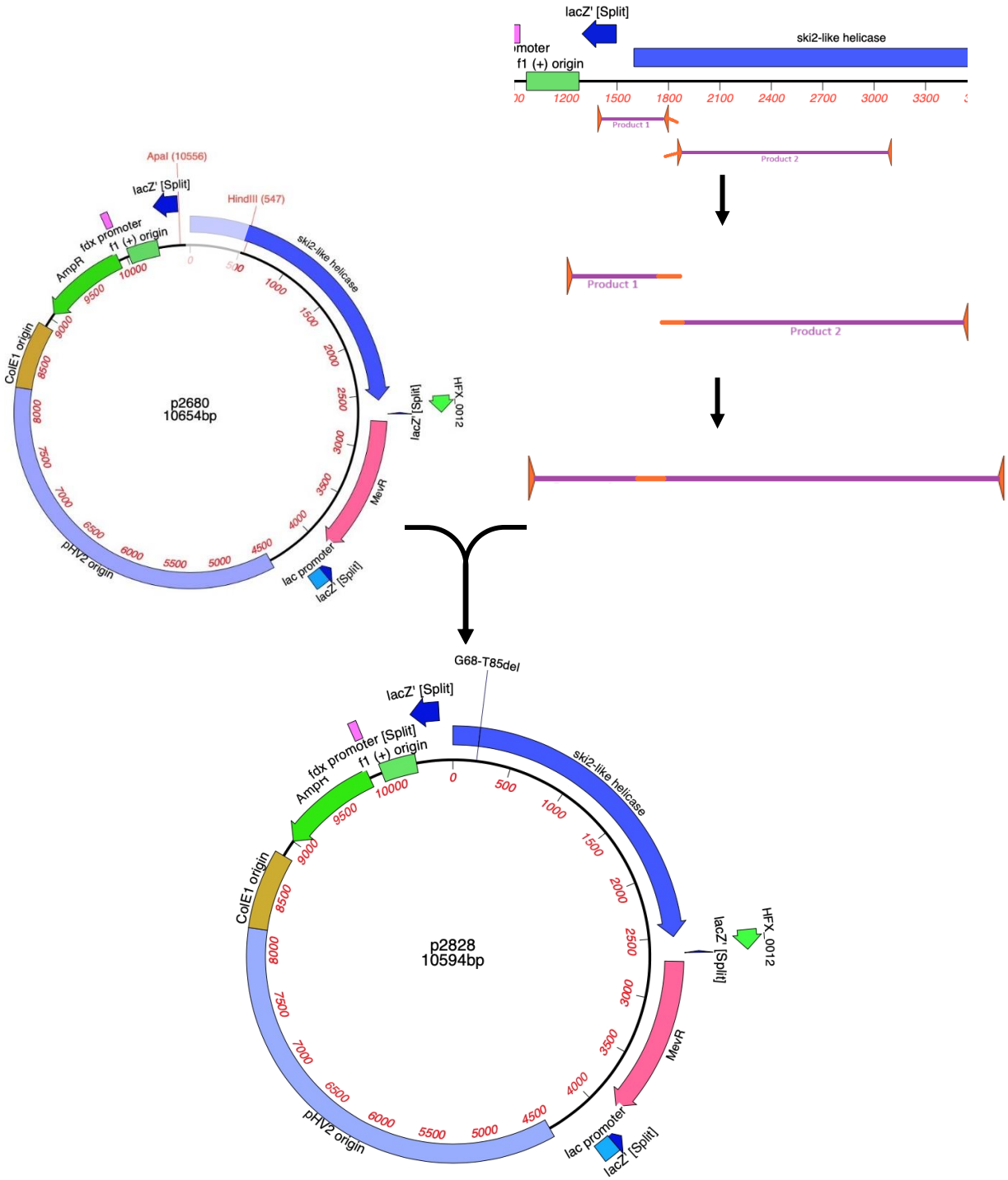


Figure 50. Generation of the p2828 plasmid. The G68-T85del mutation was produced by overlap extension, through amplification of part of the H. mediterranei *hel308 gene from p2680. PCR product was fragment-swapped onto the chromosome between the ApaI and HindIII sites.*

The *H. mediterranei hel308-G68-T85del* (corresponding to deletion of the 18 amino acid "proline loop") was produced by overlap extension PCR and then fragment swap. Overlap extension PCR was performed using four primers (o113, o867, o2472 and o2473), two of which (o2472 and o2473) have long tails with complementary sequences. Two separate PCR reactions (o113 and o2473 to generate product 1, and o867 and o2472 to generate product 2) were carried out, amplifying the sequences to either side of the sequence to be deleted. The products of these reactions were then purified and mixed, and 10 rounds of amplification carried out in the absence of the external primers. This allows

the complementary sequences introduced by the long-tailed middle primers to overlap, forming one product in which the 54 bases corresponding to G68-T85 have been deleted. The product was then cut at the ApaI and NdeI sites already present in the amplified region. Once ligated into the plasmid, the deletion was confirmed by sequencing.

5.3.1.1 Production of *dam*⁻ plasmids

As the strains used in this series of experiments are not *Δmrr*, they would not tolerate *dam*-methylated DNA well. As the usual *E. coli* strains used in preparation of plasmids are not *mrr*-deleted, the DNA produced by them will bear these methylation patterns. For this reason, plasmids are passaged in *Δmrr E. coli* strains prior to transformation into *Haloferax*, where needed. However, it was noted that transformation of the above plasmids into *dam*⁻ hosts resulted in low- or no-transformation efficiency. This has previously been reported in some other instances, with different plasmids (Smith, 2021). To examine the phenomenon in more detail, the following plasmids were used:

Table 5-1 Plasmids used in this section

Plasmid	**Size**	**Notes**
p23	2.9kb	Empty vector; control for transformation efficiency
p2680	10.4kb	Complementation vector containing *H. mediterranei hel308*
p2861	10.6kb	Complementation vector containing *H. volcanii hel308*
p2828	10.4kb	Complementation vector containing *H. mediterranei hel308-G68-T85del*

All of the above plasmids bear the *ampR* marker gene, allowing detection of their presence in the *E. coli* hosts.

To examine the qualities of the *E. coli* cause incompatibility with the plasmids, several *E. coli* variants were utilised. The full list of genomic edits of the strains are listed in Chapter 2, but pertinent details are summarised below:

Table 5-2 Summary of the E. coli *strains used in this section*

Strain	**Notes**
XL-Blue	Typical strain used for everyday cloning in the lab. *ΔrecA*
E14	*Δdam*
E15	*Δdam* *ΔmutS*
E16	*Δdam* *ΔmutS* *ΔrecA*

Like many laboratory strains of *E. coli,* the *recA* gene has been deleted from XL-Blue. RecA protein binds to ssDNA, forming a nucleoprotein filament. This filament can then seek out homologous duplex sequences, encouraging strand exchange and recombination (Chen et al., 2008). Deletion of *recA* greatly reduces the rate of recombination within the cells, meaning that plasmids carried by this strain are less likely to recombine into the host genome. It is therefore a desirable quality for laboratory cloning applications.

Dam (DNA adenine methylase) is an *E. coli* enzyme that adds methyl groups to the adenine of the sequence 5'-GATC-3' (Geier & Modrich, 1979). This modification is not inherited during DNA replication, meaning that, immediately following replication, the parent and daughter strands can be identified through the presence and absence of methylation, respectively. In cases of DNA mismatch generated during replication, it allows repair enzymes to correct the mismatch to match the parent strand sequence. However, deletion of *dam* does increase the mutation rate, as mismatches produced during replication are "corrected" at random – with equal chances of repairing the parental sequence or fixation of the mutation in both strands (Marinus, 2010).

MutS is an *E. coli* protein that plays an important role in mismatch repair. It acts as part of a complex, wherein its role is to identify mismatches and small insertion/deletions in hemimethylated DNA following replication. It then recruits MutL and the endonuclease MutH. MutH nicks the DNA backbone of the unmethylated strand, as its activity is inhibited by methylation. In strains with wild-type methylation patterns, this mismatch repair complex is a useful tool to prevent repair through recombination (Calmann & Marinus, 2004).

Deletion of *dam* is desirable to produce the unmethylated DNA preferred by *Haloferax*, which cleaves methylated DNA via the Mrr restriction enzyme. However, deletion of *dam* from *E. coli* strains results in increased mutability and recombination, as mismatch repair complexes then lack directionality with regard to parent/daughter strands. These complexes cause nicks in the backbone at any unmethylated 5'-GATC-3' sequences, resulting in double-strand breaks that must then be repaired through recombination. While deletion of *recA* is also desirable to improve stability of plasmids in *Δdam* strains, this is not possible, as it removes the mechanism by which the many double-stand breaks in this background would be repaired. Double deletion of *dam* and *recA* has been shown to be lethal in *E. coli* (Marinus, 2000).

The solution to allow production of good-quality unmethylated DNA is therefore to delete all three of the above-mentioned genes. Deletion of *mutS* allows deletion of *dam* and *recA* to be tolerated together, as it reduces the amount of double-strand breaks that would otherwise need to be repaired through recombination. However, the combination of these three alterations does result in an increased mutation rate, hence why the XL-Blue strain is still used for day-to-day cloning applications, and the *Δdam, ΔmutS ΔrecA* strain used only for brief passage of plasmids prior to transformation into *Haloferax*.

While transformation efficiency is typically reported in terms of the nanograms of DNA used, this would not be appropriate in this instance given the difference in size between the plasmids used. To remedy this, plasmid concentration was adjusted relative to the size of each, to achieve a similar molarity of plasmid in each case.

Following transformation into the *E. coli* strains by electroporation, transformants were plated onto LB plates containing ampicillin. Both an undiluted sample, and a 10^{-2} dilution were plated for each combination of plasmid and strain. The plates were placed in a 37°C incubator overnight. The following day, colonies were counted, yielding the results below. In the case of p23, all of the *E. coli* strains produced a lawn of colonies, which could not be counted. For these combinations, the plates from the 10^{-2} dilution were used, and the results multiplied by 100 to give an approximate value. The data is shown in Figure 51, below.

	p23	p2680	p2861	p2828
E14	103,200	1	8	8
E15	162,500	4	40	74
E16	79,200	2	1	13
XL-Blue	240,000	2372	2868	2448

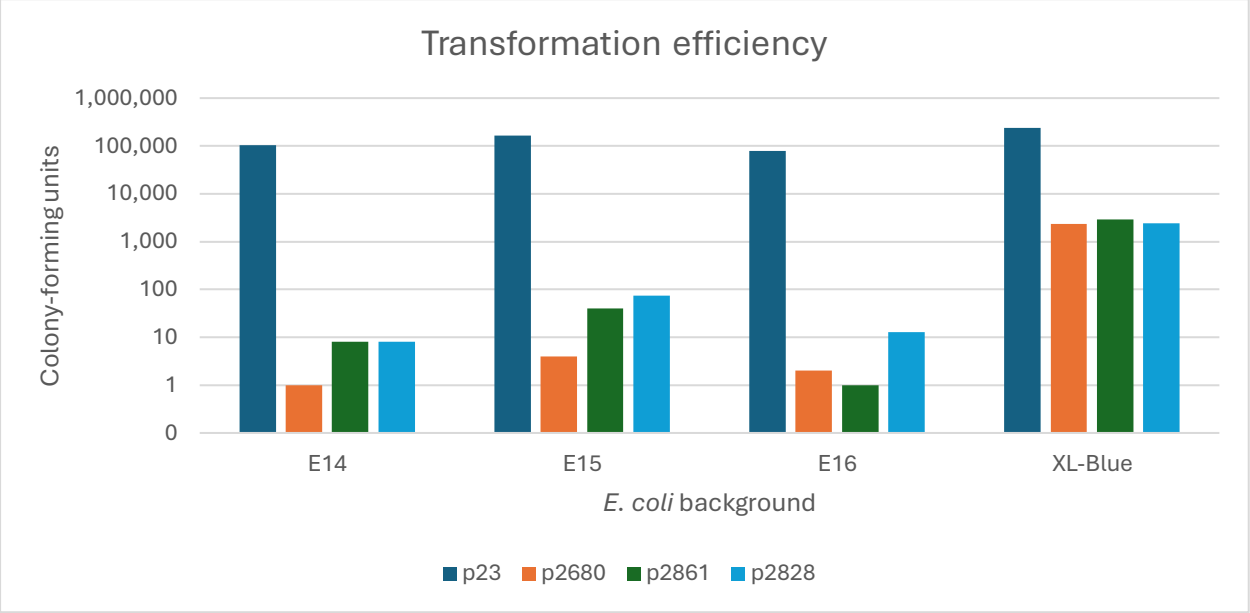


Figure 51. Transformation efficiencies of the four plasmids. Note that the y axis is logarithmic, as all bars other than those corresponding to p23 were rendered invisible when using a linear axis.

Significant differences can be seen between both the plasmids and the *E. coli* strains. However, it is difficult to ascertain whether some part of the difference between the *hel308*-containing plasmids and p23 could be due to quality of the plasmid preparation, as the transformation efficiency in XL-Blue is still almost 100-fold lower than that of the control plasmid. It is also possible that some of the variation in transformation efficiency between the p23 control and the experimental plasmids is due to the sheer size of the larger plasmids. There are also some very large discrepancies between the transformation efficiencies of p23 into the various *E. coli* strains.

However, it seems clear that the other *E. coli* strains tolerate the *hel308*-containing plasmids much worse than the everyday cloning strain. This could indicate that the genes in question are both expressed by the host, and are able to fold sufficiently in the cells' internal environment to remain at least somewhat active. This is unexpected, as proteins from halophilic organisms frequently misfold and become inactive under low-salt conditions (Arakawa et al., 2017). However, it is also possible that the differences observed here are due to the size difference between the experimental and control plasmid, or may be caused by another of the genes present on these larger plasmids.

It is also not known what particular quality allows the survival of the few E14, E15 and E16 cells that do survive with the *hel308*-bearing plasmids. It is possible that the stable transformation of this strain-plasmid combination is only possible as a result of a random mutation affecting the expression or activity of the *hel308* protein within these cells.

5.3.2 Strain construction

Due to difficulties in producing dam$^-$ forms of the above plasmids, the dam$^+$ forms of each were used instead. While this significantly affects the transformation efficiency into *Haloferax*, no alternative

was available, as no *Δmrr* strain of *H. mediterranei* has yet been produced. Presence of the episome was confirmed by colony PCR.

Table 5-3 Strains used in this section

Strain	Genotype	Notes
H5222	*ΔpyrE2* *ΔtrpA* *hel308::[Δhel308::trpA+ pyrE2+]*	Produced by A. Dattani Pop-in of a *trp*-marked *hel308* deletion vector (p2576)
H5580	*ΔpyrE2* *ΔtrpA* *hel308::[Δhel308::trpA+ pyrE2+]* *[Hmed hel308 + MevR+]*	Produced for this study H5222 with p2680 complementation (*H. mediterranei hel308)*
H5594	*ΔpyrE2* *ΔtrpA* *hel308::[Δhel308::trpA+ pyrE2+]* *[Hmed hel308-G68-T85del+ MevR+]*	Produced for this study H5222 with p2829 complementation (*H. mediterranei hel308-G68-T85del)*
H5612	*ΔpyrE2* *ΔtrpA* *hel308::[Δhel308::trpA+ pyrE2+]* *[Hmed hel308 + MevR+]*	Produced for this study H5222 with p2862 complementation (*H. volcanii hel308)*

5.3.3 Results

Following preparation of the strains above, deletion of chromosomal *hel308* was attempted to see in which circumstances this was tolerated. H5222 (no complementation), H5580 (*H. mediterranei hel308* complementation), H5594 (*H. mediterranei hel308-G68-T85del* complementation) and H5612 (*H. volcanii hel308* complementation) were all grown in minimal-phosphate media with added uracil for 5 successive 5ml overnights. Under these low phosphate conditions, ploidy of *Haloferax* reduces drastically (Zerulla et al., 2014), increasing the chances of isolating strains homozygous for the desired deletion. Following this, cells were plated on Cas-based media containing 5'FOA and uracil. This would select against presence of the *pyrE2*-marked pop-in vector, but would not select against the MevR-marked complementing episome. Colonies growing on the 5'FOA plates were then restreaked onto Cas-based plates supplemented with uracil, to reduce exposure to the mutagenic 5'FOA (Wellington & Rustchenko, 2005).

As the deletion vector is marked with the *trpA* gene, deletion of *hel308* should result in *trpA* at the *hel308* locus; the absence of tryptophan in the media is therefore a selector for *hel308* deletion. However, *H. mediterranei* is highly polyploid, meaning that further steps must be taken to ensure that all copies of *hel308* have been deleted from the genome.

Colonies could not be tested for *hel308* deletion by colony hyb, as the presence of this gene in the episome would cause hybridisation even in the absence of the chromosomal copy of the gene. Instead, candidates were tested by colony PCR, with the intention to confirm promising candidates through Southern blot.

Colony PCRs were carried out using the primers o306 and o2264, as shown in Figure 52, below. Of these, one anneals within the *hel308* gene, while the other binds upstream on the chromosome; as a result, this should not cause amplification from the complementing episome. H828 was included as a positive control.

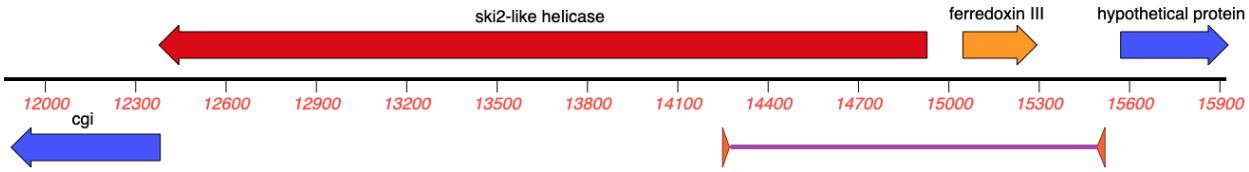


Figure 52. Position of primer binding for identification of Δhel308 *candidates by colony PCR.*

However, no pop-outs were isolated from the above process, even in H5580, which should have been fully complemented with wild-type *hel308.* Closer examination of p2576, the deletion vector used to produce H5222, revealed that the chromosomal section targeted for deletion by this vector was slightly longer than the length of *H. mediterranei hel308*. This results in deletion of the start codon of the *cgi* gene, which overlaps with the end of the *hel308* gene, as is the case in *H. volcanii*. This is shown in Figure 53, below. This was concerning, as this gene is essential in many species due to its role in the KEOPS complex (Naor et al., 2012). As p2576 was also used in the experiment that initially identified the essentiality of *hel308* in *H. mediterranei* (Dattani, unpublished research), it is possible that it was disruption of the *cgi* gene instead that made deletion non-viable.

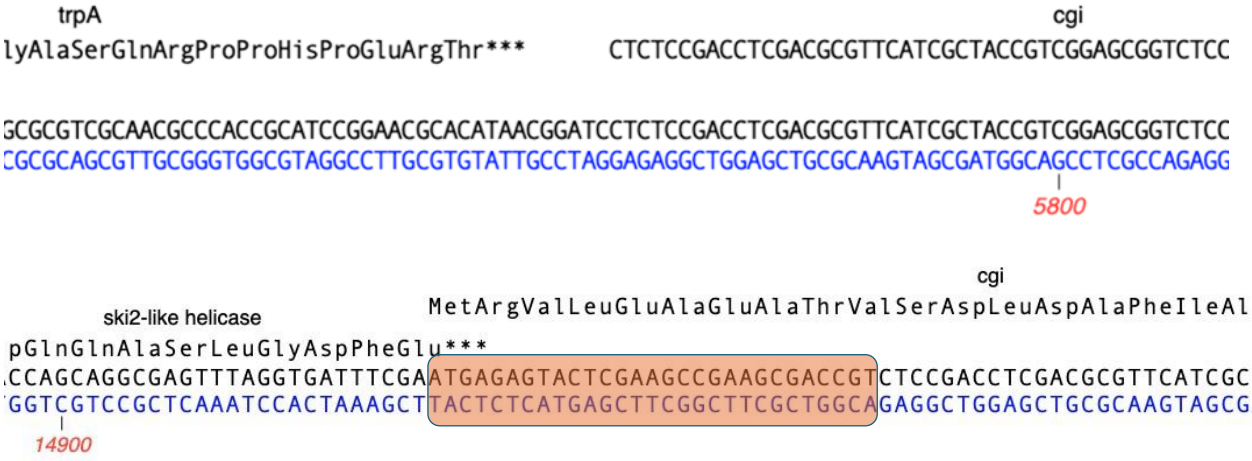


Figure 53. DNA sequence of the start of the cgi *gene in the p2576 deletion vector (above) and the wild-type chromosomal sequence (below). Bases inadvertently deleted in the deletion vector are highlighted in red. This includes the start codon.*

While it is possible that the *cgi* gene is not essential in *H. mediterranei*, or that *hel308* is essential, the deletion of both genes in the deletion vector used to determine *hel308* essentiality was considered a very significant confounding factor. In light of this, the complementation experiment was abandoned until the essentiality of *hel308* could be verified.

5.4 Attempted deletion of *H. mediterranei hel308*

5.4.1 Plasmid construction

In order to determine the essentiality of *hel308 and cgi* in *H. mediterranei*, three new deletion vectors were designed. These would attempt to delete the *hel308* and *cgi* genes individually, as well as together. If the *cgi* gene is essential, then only the *hel308* deletion would be possible.

Primers were designed to amplify the relevant upstream and downstream sequences of the two genes. Their binding positions on the chromosome and the restriction sites that they introduce are shown in Figure 54, below:

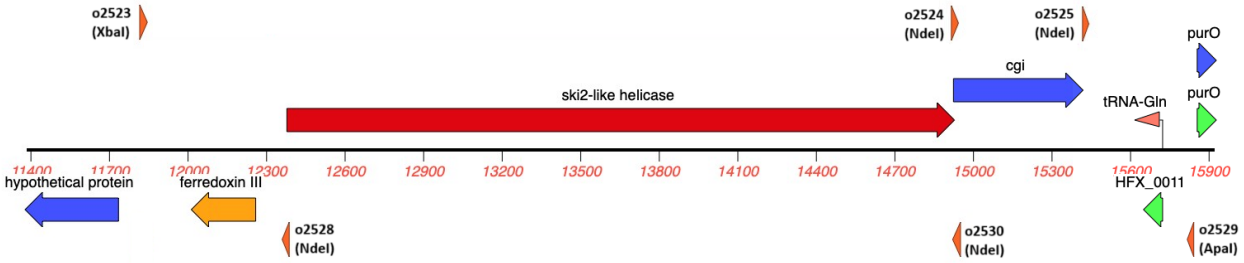


Figure 54. Annealing sites of primers designed for construction of H. mediterranei hel308 *and* cgi *deletion vectors.*

These primers were then used in separate reactions to amplify the relevant homology arms upstream and downstream for each gene combination. This was performed for each of the three deletion vectors, as shown in Figures 55-57, below:

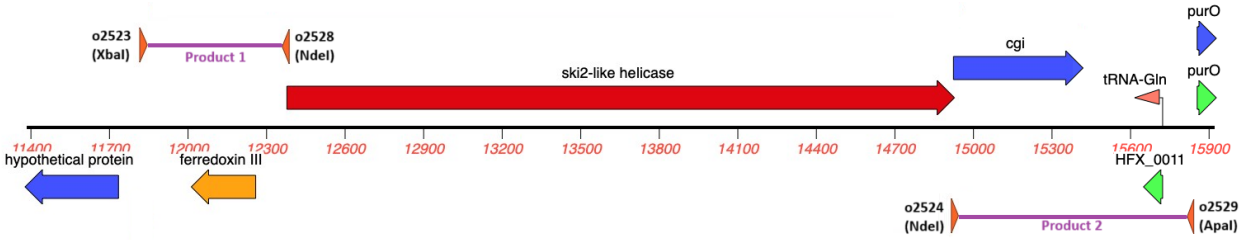


Figure 55. Regions amplified for construction of the H. mediterranei hel308 *deletion vector.*

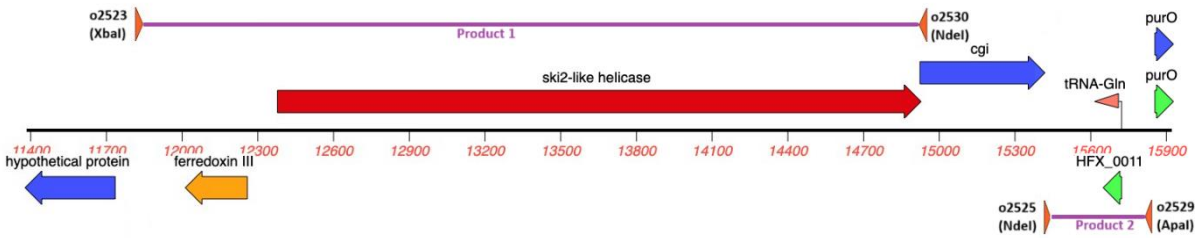


Figure 56. Regions amplified for construction of the H. mediterranei cgi *deletion vector.*

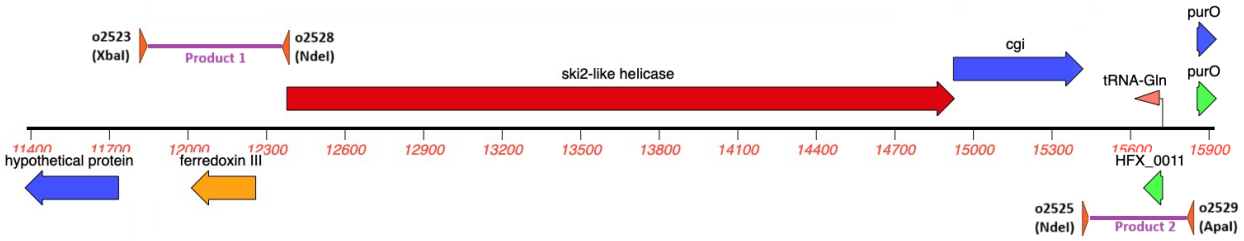


Figure 57. Regions amplified for construction of the H. mediterranei hel308 and cgi *deletion vector.*

The PCR products were then cut at the NdeI site and ligated together, uniting the upstream and downstream sections with the NdeI site in the middle of the sequence. These were then each inserted into vector p131, which carries the *pyrE2* marker, but does not carry a *Haloferax*-compatible origin. Diagrams of the resulting plasmids are shown in Figure 58, below:

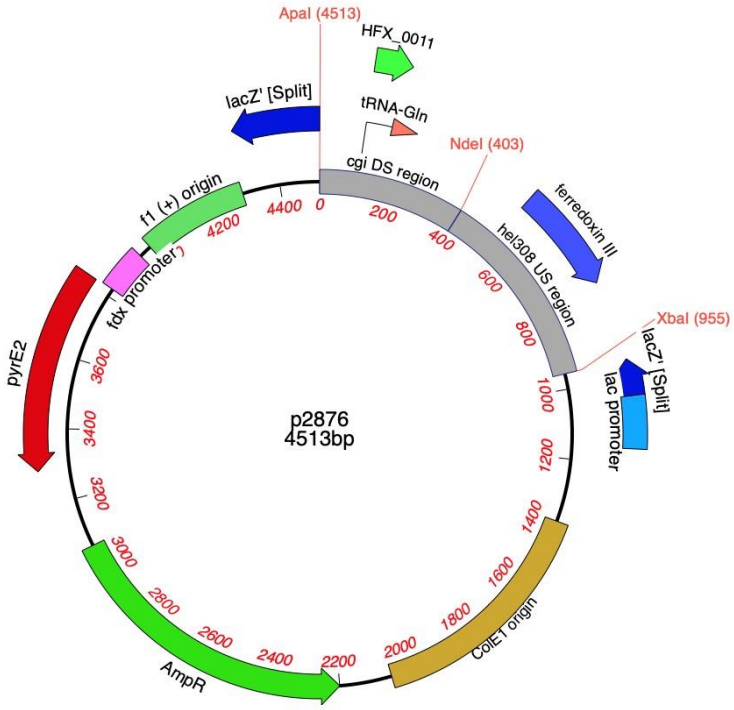


Figure 58. p2876, vector for deletion of H. mediterranei hel308 and cgi genes.

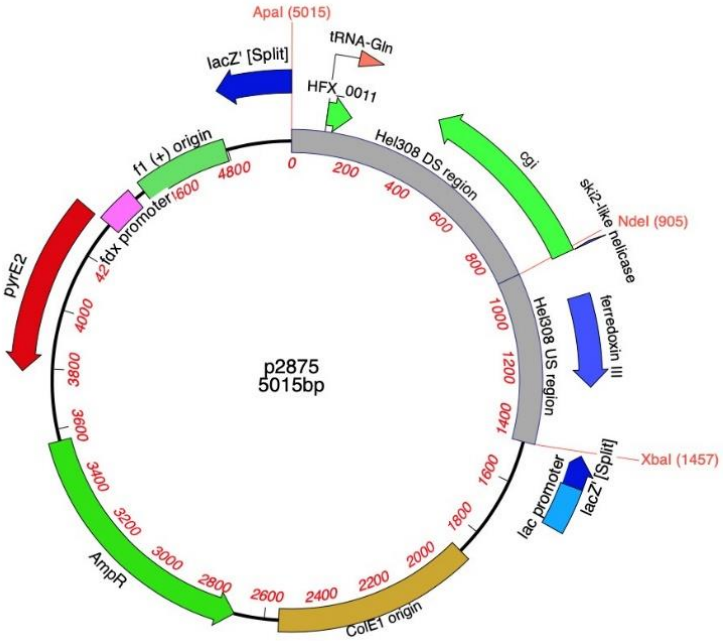


Figure 59. p2875, vector for deletion of H. mediterranei hel308.

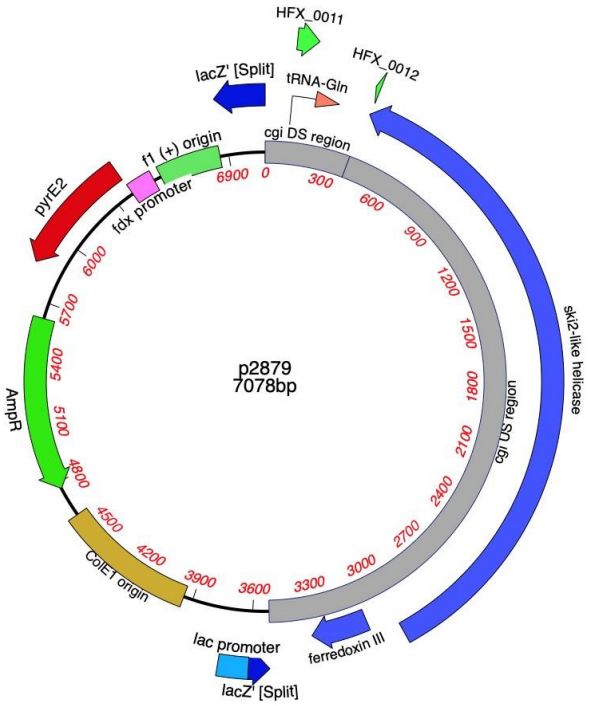


Figure 60. p2879, vector for deletion of H. mediterranei cgi.

In *H. volcanii,* the *hel308* gene is somewhat difficult to delete, as the growth defect that is incurred by its deletion effectively applies a selective pressure against cells that have lost this gene (T. Allers, personal communication). The *trpA* gene was thus added to each of the three deletion vectors at the NdeI site between the two homology arms. In pop-outs that have successfully deleted the gene/s (as opposed to reverting to wild-type), this marker will be present on the chromosome, replacing the gene/s. While this does not solve the complicating factor of the high ploidy (and thus high chance of merodiploidy in pop-out candidates), this additional selective marker should allow confidence that ura- trp+ candidates contain at least one chromosome bearing the deletion.

The *trpA* gene was excised from p1277, where an NdeI site is present at each end. It was then ligated into each of the three deletion vectors, following their digestion with NdeI and treatment with rSAP, to prevent re-ligation of the compatible ends of the vector. The resulting plasmids are shown in Figures 61-63, below.

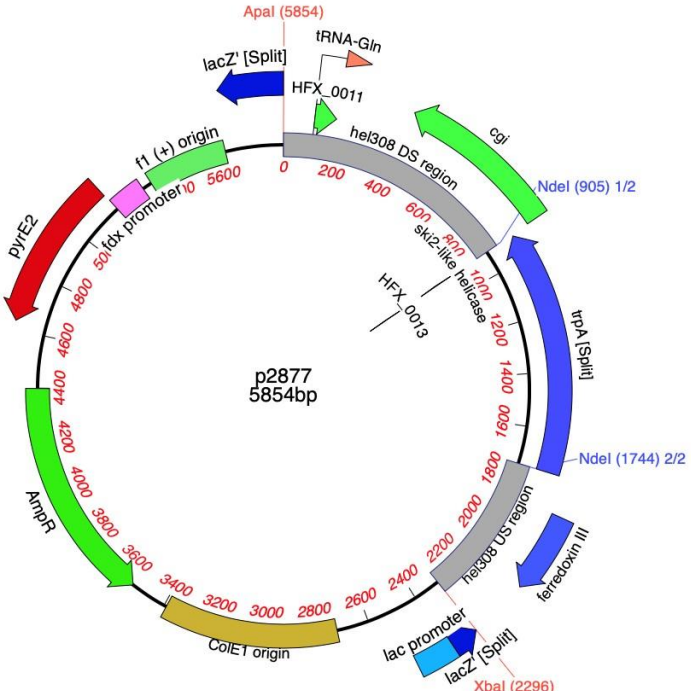


Figure 61. p2877, trpA-*marked vector for deletion of* H. mediterranei hel308 *gene.*

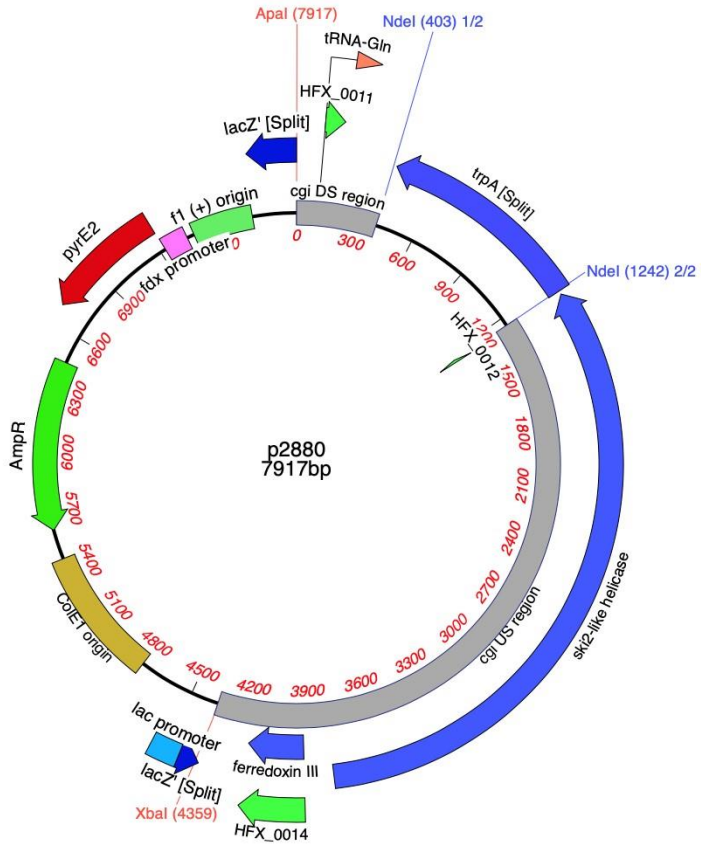


Figure 62. p2880, trpA-*marked vector for deletion of* H. mediterranei *cgi genes.*

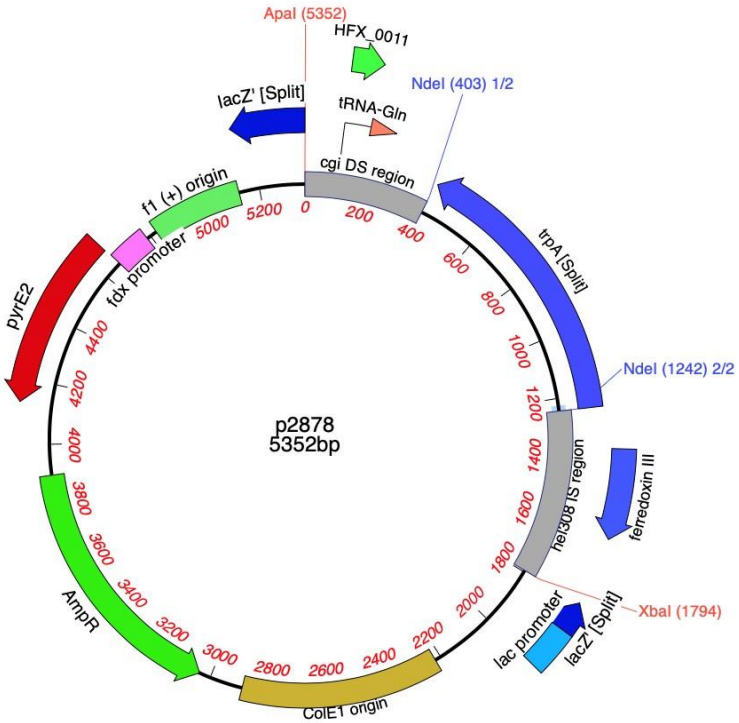


Figure 63. p2878, trpA-*marked vector for deletion of* H. mediterranei hel308 *and* cgi *genes.*

Note that the new *H. mediterranei hel308* deletion vector (p2877) does not disrupt the *cgi* gene, which would be left intact following deletion, as shown in figure 64, below.

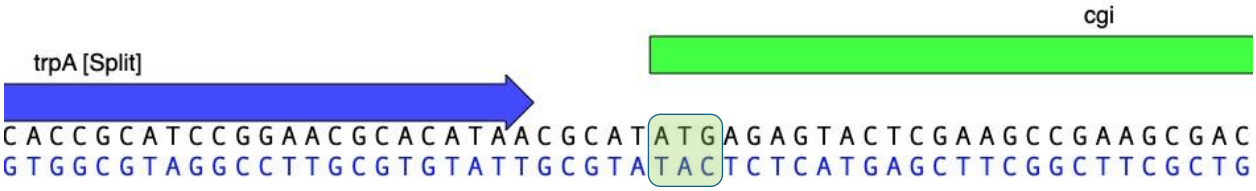


Figure 64. Start of the cgi *gene in p2877. Start codon is indicated by the green box.*

In each case, presence of the *trpA* gene (and its orientation in line with the deleted gene/s) was confirmed by colony PCR, to ensure that it would be inserted in the same orientation of the original gene. All three vectors were passaged in *dam*$^{-}$ *E. coli* strains to remove methylation patterns. This was necessary, as no *H. mediterranei Δmrr* strains have yet been created in our lab. The new nomenclature for the plasmids is shown below.

Table 5-4 Plasmids used in this section

***dam+* plasmid number**	***dam*- plasmid number**	**Purpose**
p2877	p2881	*hel308* deletion vector
p2878	p2882	*cgi* deletion vector
p2880	p2883	*hel308* and *cgi* deletion vector

5.4.2 Strain construction

Deletion vectors were transformed into H828 (Δ*pyrE2, ΔtrpA*) and were confirmed by colony PCR.

Out of curiosity, p2881, the *hel308* deletion vector, was also transformed into H4676 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*), with the eventual aim of replacing the *H. mediterranei hel308* gene with the *H. volcanii hel308* gene in this background, to observe whether this allowed deletion of the final origin. This produced the following strains:

Table 5-4 Strains used in this section

Strain identifier	**Genotype**	**Notes**
H5739	*ΔpyrE2* *ΔtrpA* *hel308+::[Δhel308::trpA+ pyrE2+]*	H828 + p2881
H5740	*ΔpyrE2* *ΔtrpA* *cgi+::[Δcgi::trpA+ pyrE2+]*	H828 + p2882
H5741	*ΔpyrE2* *ΔtrpA* *hel308+::[Δhel308 Δcgi::trpA+ pyrE2+]*	H828 + p2883
H5742	*ΔpyrE2* *ΔtrpA* *ΔoriC1*	H4676 + p2881

	ΔoriC2 *ΔoriC3* *hel308+::[Δhel308::trpA+ pyrE2+]*	

5.4.3 Results

Pop-outs of the integrated plasmids were attempted by inoculating a small (5ml) culture of Hv-min+ura with one verified colony, and incubating overnight in a 45°C rotating incubator. Each day for five days, a 20µl sample from the overnight culture was used to inoculate a fresh 5ml culture of Hv-min+ura, until five successive cultures had been produced. At the end of the five days, 1ml of the cultures were spun down and resuspended in 1ml of 18% salt water. 100µl of this culture (and a 10^{-2} dilution) was spread on Cas+5'FOA plates. This resuspension in salt water was necessary to remove uracil from the solution, which could otherwise interfere with the 5'FOA selection. As Cas+5'FOA contains no tryptophan, cells bearing no chromosomes in which the gene/s of interest had been replaced with the *trpA* marker would not be able to grow.

Unlike the *H. volcanii* form of this protocol, Hv-min broth is used for the overnight cultures of *H. mediterranei* pop-outs, as the low phosphate levels cause correspondingly low ploidy in *Haloferax* (Zerulla et al., 2014). It was hoped that this would increase the chances of isolating a homoploid colony bearing the desired deletion.

Following growth of the colonies on the selective plates, individual colonies were restreaked onto Cas+ura plates with a toothpick, to reduce the amount of time that colonies spent on the mutagenic 5'FOA plates (Wellington & Rustchenko, 2005). H828 was also streaked onto the plates as a positive control, and to allow easy identification of candidates by providing a distinguishing mark on colony lifts. H1391 was also included on the plates as a negative control in the case of Δhel308 and Δhel308Δcgi candidates, as this *H. volcanii* strain lacks the *hel308* gene and would therefore not hybridise to probes specific to this gene.

Colony lifts were then performed, and the membranes prepared for hybridisation. Probes corresponding to the *hel308* gene were used for H5611, H5739 and H5741, and probes corresponding to the *cgi* gene used for H5740. Probing for *hel308* was considered beneficial in the cases of *Δhel308Δcgi* candidates, as this allows inclusion of a negative control from *H. volcanii*. *cgi*, however, has not been successfully deleted from either *Haloferax* species, and so no negative controls were available. Results of the colony hybridisations are shown in Figures 65 and 66, below.

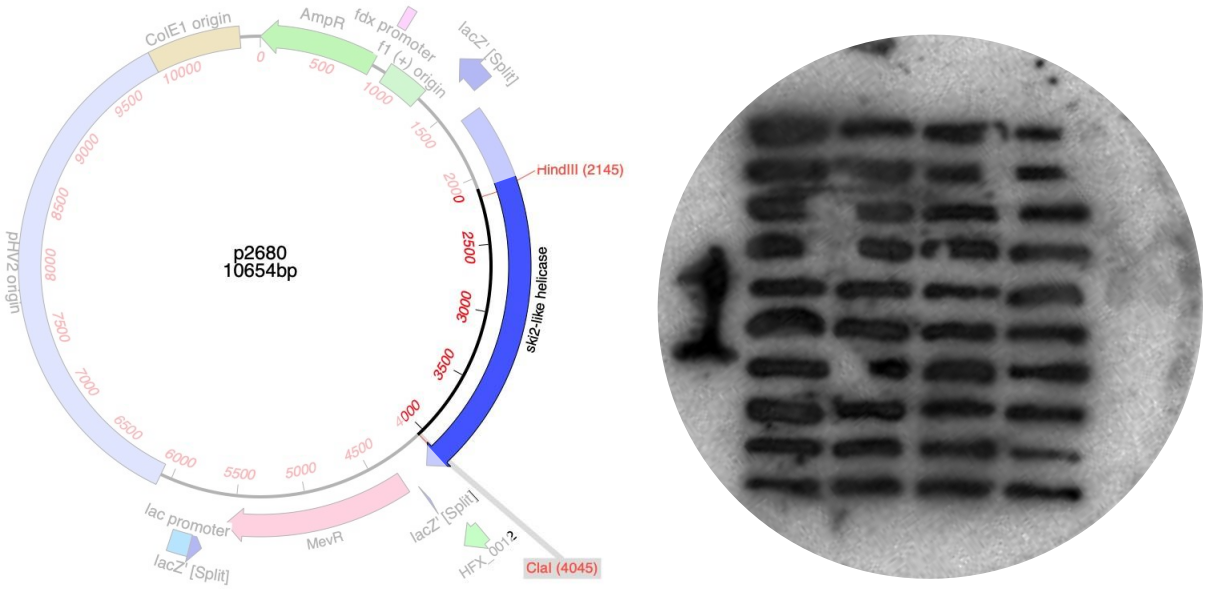


Figure 65. Colony hybridisation of H5739. Left, digestion of p2680, producing H. mediterranei hel308 *probe. Right, membrane after colony hybridisation, imaged as per the methods chapter. Positive control (H828) is present as a numeral (left side). Negative control (H1391) is present as a faint cross (right side).*

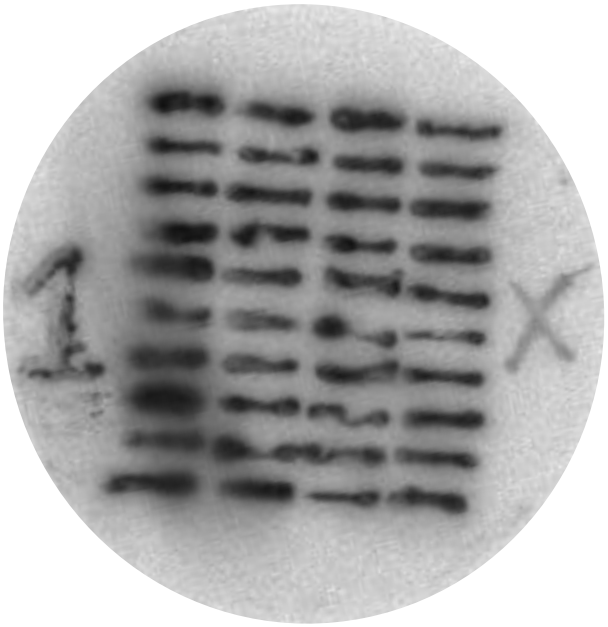


Figure 66. Colony hybridisation of H5741. Probe was produced as per Figure 65. Positive control (H828) is present as a numeral (left). Negative control (H1391) is present as a faint cross (right).

For H5740 (*Δcgi* pop-in), a hydridisation probe was created by amplification of the *cgi* region with o2524 and o2529 (shown in figure 67, below). This PCR product was then digested with PstI and MluI.

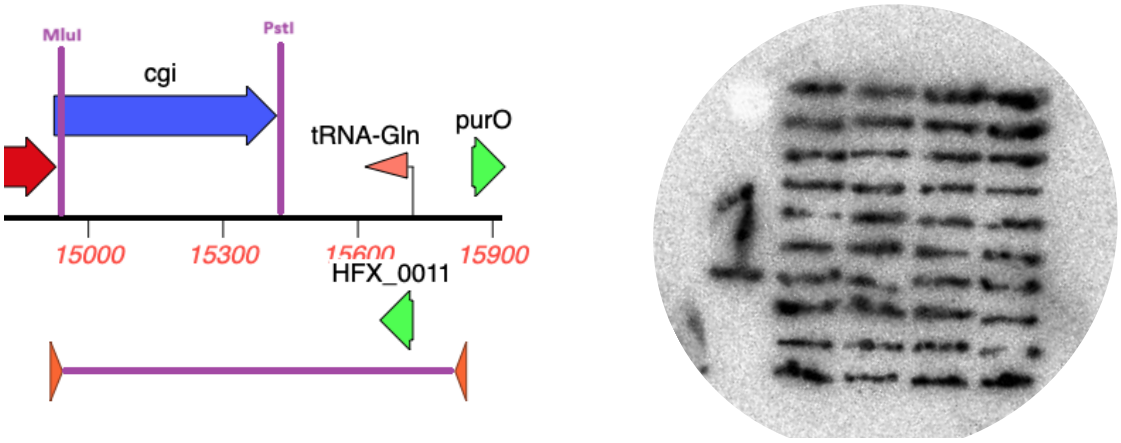


Figure 67 Production of probe for H. mediterranei cgi *gene (left) and colony hybridisation using that probe (right). H828* mrr+ *control is present as a numeral on the plate.*

Deletion of *cgi* was not expected to be viable, so attempts to produce deletion candidates from H5740 and H5741 were discontinued after a patch plate of 40 colonies each. However, it had been hoped that deletion of *hel308* may prove possible for H5739. However, no candidates were identified following screening of 400 candidates. Time restrictions precluded pursuing this goal further.

5.5 Attempted deletion of *H. mediterranei mrr*

Following the difficulty of constructing *dam*- forms of some of the plasmids (as described in chapter 5.3.1, above), it appeared that construction of a *Δmrr H. mediterranei* strain would be beneficial, in case such issues arose again. Deletion of this gene is commonplace in *H. volcanii*, but had not previously been performed in *H. mediterranei* in our lab.

5.5.1 Strain construction

A vector for deletion of *H. mediterranei mrr* had previously been constructed by A. Dattani, but had not been successfully used. This vector is shown in Figure 68, below.

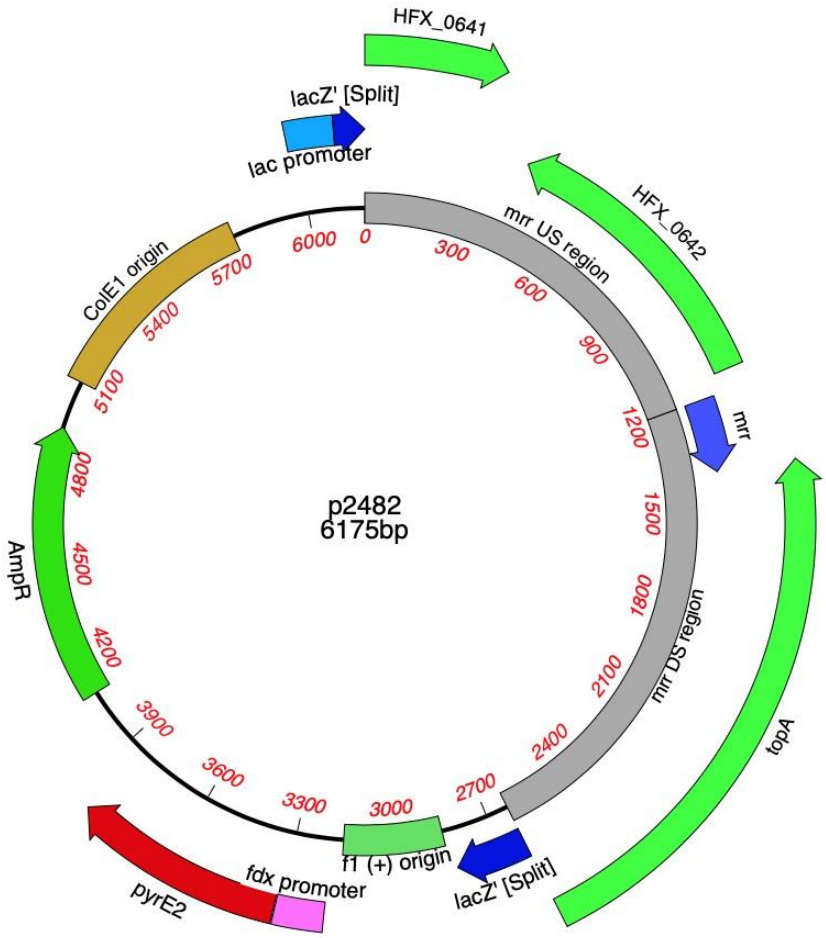


Figure 68. p2482, deletion vector for H. mediterranei mrr *gene. The dam- preparation of this plasmid is p2483.*

As can be seen in figure 68, this deletion vector is slightly different to previous deletion vectors discussed, as around 200 bases of the *mrr* gene are still present. This was a deliberate choice on the part of A. Dattani to ensure that that *topA* gene (encoding topoisomerase A) was not disrupted. The gene annotated *mrr* in the above diagram is in fact only a small part of the 1005-base long gene; the start codon and majority of the length of the gene are deleted. In addition, although it appears close to the junction between the two homology arms, gene HXF_0642 (encoding a predicted geranylgeranylglyceryl phosphate synthase) is not disrupted by this vector, as shown in Figure 69, below.

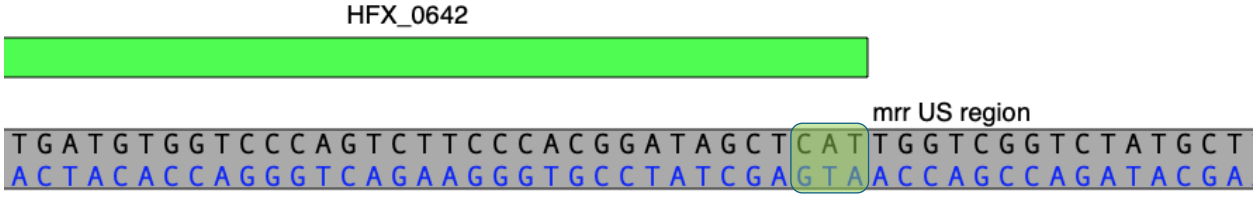


Figure 69. The start of the predicted geranylgeranylglyceryl phosphate synthase gene in p2482. The start codon is not disrupted, and is highlighted in green.

p2483 (the *dam*- preparation of p2482) was transformed into H828 to produce pop-in H5682. Pop-in was confirmed by colony PCR.

5.5.2 Results

Following confirmation of pop-in by colony PCR, pop-outs were generated, as described above. These were then subjected to colony lifts in preparation for examination by colony hybridisation. The probe for colony hybridisation was generated by amplification of part of the *mrr* gene from H828 using primers o2535 and o2536. Annealing sites for these primers are shown in figure 70, below.

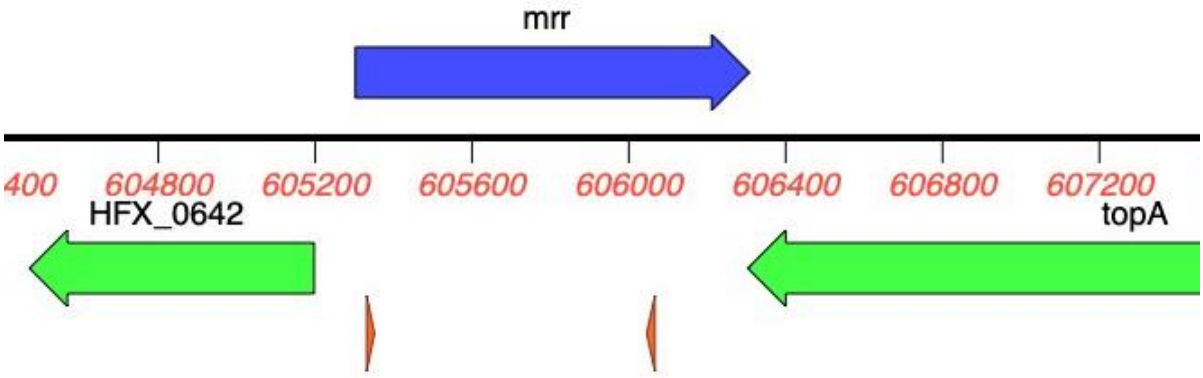


Figure 70. Annealing sites of primers o2535 (left) and o2536 (right).

Radiolabelled DNA probe was generated from the PCR product, and colony hybridisation performed as per the methods chapter. Results are shown in Figure 71, below.

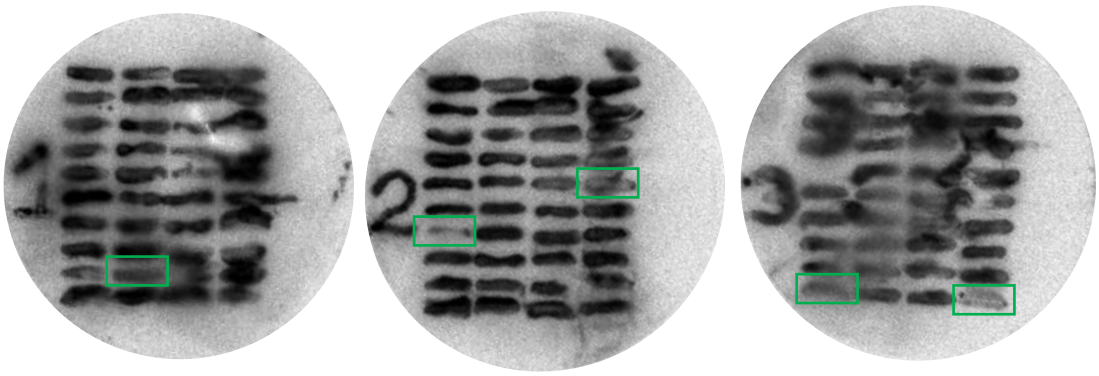


Figure 71. Colony hybridisation of 120 Δmrr *candidates. Promising colonies are indicated by the green boxes.*

Genomic DNA was prepared from the five promising candidates (and H828 as a *mrr+* control) and digested with ApaLI and NruI-HF. This is predicted to produce a 3kb band if *mrr* is present, and a 2.2kb band if the gene has been successfully deleted. Diagram of the *mrr* locus showing the restriction sites is shown in Figure 72, below.

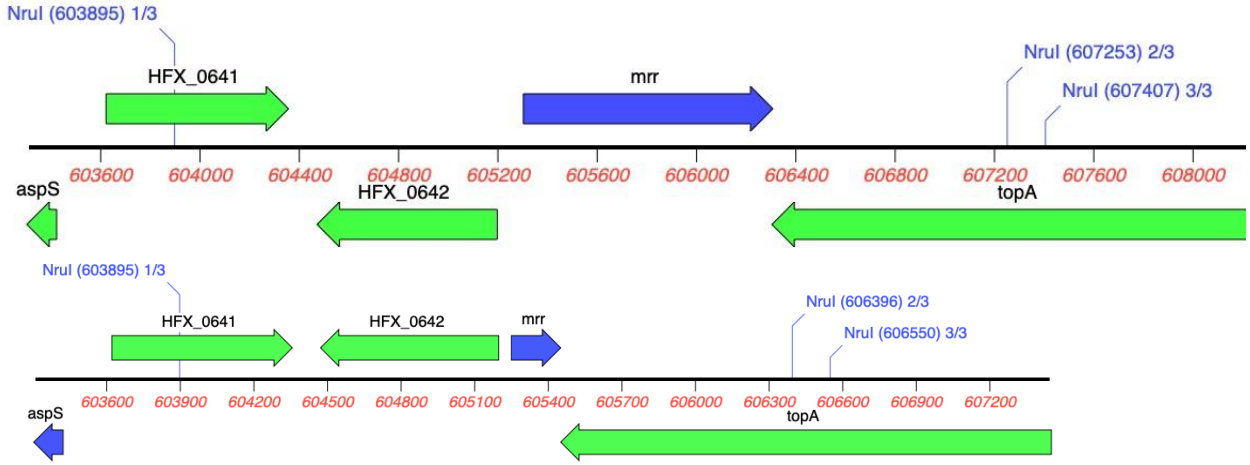


Figure 72. Diagram of the H. mediterranei mrr *locus before deletion of* mrr *(top) and after deletion of* mrr *(bottom), showing the restriction sites used in the Southern blot.*

NruI could not easily be used to digest the deletion vector (p2482), as this restriction site is present multiple times on the plasmid and would produce several fragments of similar size. Identifying the correct fragments would be difficult. Therefore, the probe for the Southern blot was prepared from p2482, digested at the ScaI and BseYI sites. The position of the restriction sites is shown in Figure 73, below.

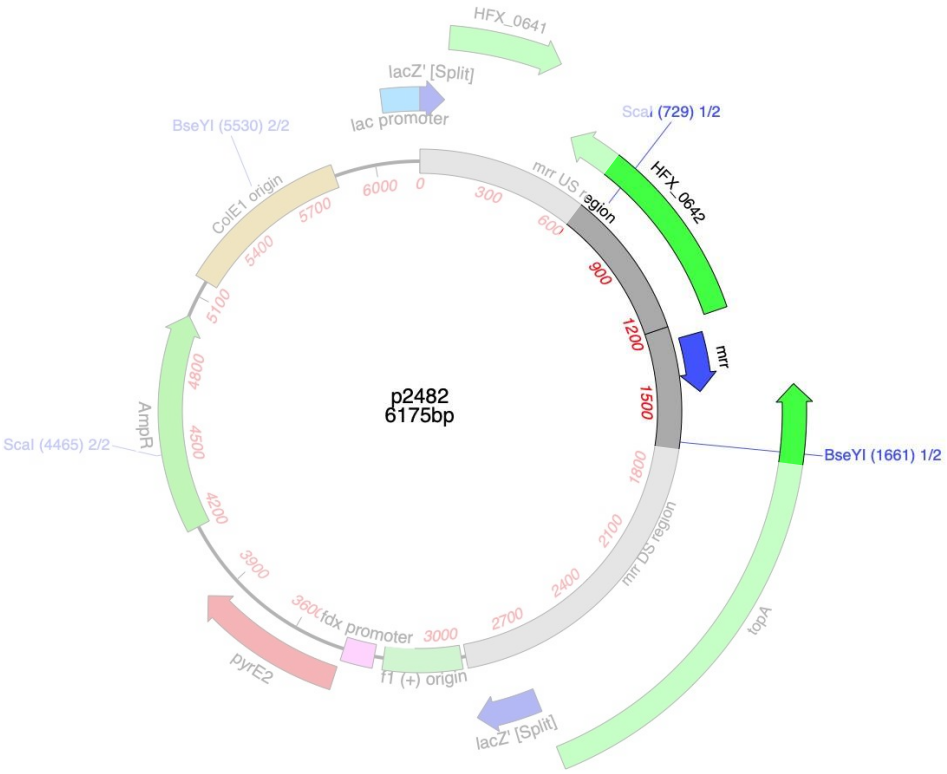


Figure 73. Digestion of p2482 used to generate the probe to identify Δmrr *genotype in* H. mediterranei. *Fragment used was that bridging the junction between the homology arms.*

Southern blot preparation, hybridisation and visualisation were performed as per the methodology chapter. Even though the colony hybridisation appeared promising, Southern blot did not confirm the presence of *Δmrr* colonies. The results are shown in Figure 74, below.

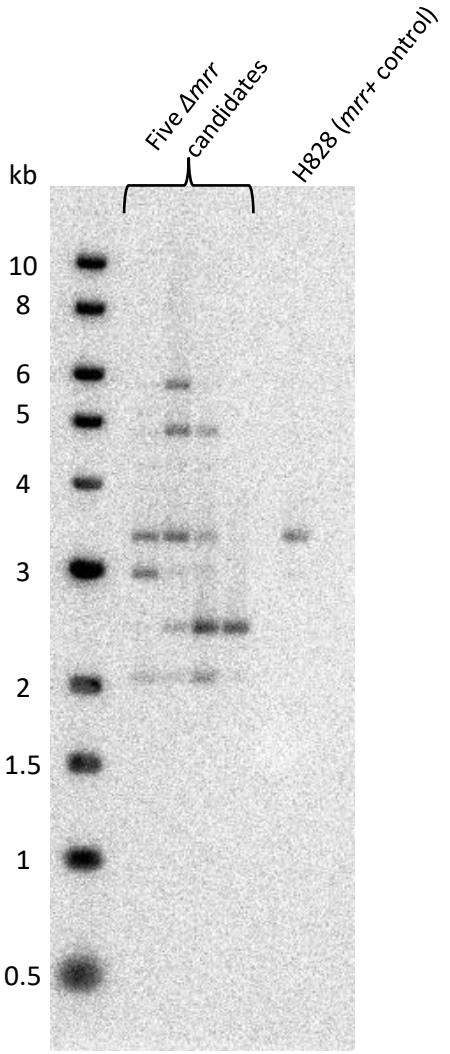


Figure 74. Southern blot for H. mediterranei Δmrr *candidates, showing unexpected banding pattern.*

Many unexpected bands were observed. While some of these could charitably be suggested to be the 2.2 and 3kb bands that the digest was expected to produce, several others were present that were not identified. This could be due to incomplete digestion; however, there are some bands below the smallest expected size (around 2kb). It is possible that the probe used hybridised to other similar regions of the genome, or that an unexpected recombination event had somehow occurred at the *mrr* locus, creating the many-banded pattern. Time restrictions precluded investigating these results further, and ultimately no *Δmrr* strains resulted from this experiment.

5.6 Future work

While the *H. mediterranei hel308* and *mrr* genes have certainly proven "difficult to delete", the essentiality of these genes are not yet conclusively determined. Several avenues are available to explore this topic further, discussed in the sections below. Should *H. mediterranei hel308* be proven to be essential through any of these experiments, the complementation assay originally described at the start of this chapter could then be carried out to ascertain what qualities of this gene are essential in this species.

5.6.1 Complementation of genes of interest

Deletion of the *hel308* and *cgi* genes could be re-attempted accompanied by complementation with the same genes on a *mevR*-marked episome. Several variants of these are already available (discussed in section 5.3.1, above). For exploration of the essentiality of *cgi* in this species, equivalent complementation plasmids could be produced bearing this gene. Pop-out of the genes from the chromosome could then be attempted, and candidates investigated via the colony PCR method described previously.

Following confirmation of deletion of the gene of interest from the chromosome, the strains can be passaged again in sequential non-selective overnights and plated onto non-selective plates. Under such conditions, if the gene of interest is inessential, some cells may lose the complementing episome. Loss of the episome can be determined by restreaking candidates from the non-selective plates onto plates containing mevinolin (or other selective plates, depending on marker genes present in the episome). While the *hrdB* marker is more commonly used for this purpose (Lestini et al., 2010), the advantage of using *mevR* in this context is that it does not require any genetic modification of the host strain prior to use. This is beneficial in *H. mediterranei,* where the higher ploidy makes isolation of homoploid colonies more difficult.

If the *hel308, cgi and mrr* genes could not be deleted in this study due to growth defects in the deletion mutants, it is possible that complementation of these genes via an episomal gene copy could allow this effect to be bypassed and homoploid deletion mutants to be successfully generated.

5.6.2 Improve efficiency of pop-in/pop-out

The pop-in/pop-out method (Bitan-Banin et al., 2003) is reliant on two instances of recombination at two different positions; either an upstream pop-in followed by a downstream pop-out, or a downstream pop-in followed by an upstream pop-out. In *H. volcanii*, this appears to occur with relative frequency, but *H. mediterranei* demonstrates significantly lower recombination levels (Dattani et al., 2022). It may thus be reasonably assumed that both pop-in and pop-out events occur with less frequency in *H. mediterranei*.

Jones (2019) previously showed that, in *H. volcanii*, frequency of recombination between two sequences is correlated with the length of the homologous sequence that they share. The relationship between homology length and recombination is exponential, where longer regions of homology produce much higher levels of recombination, as shown in Figure 75, below.

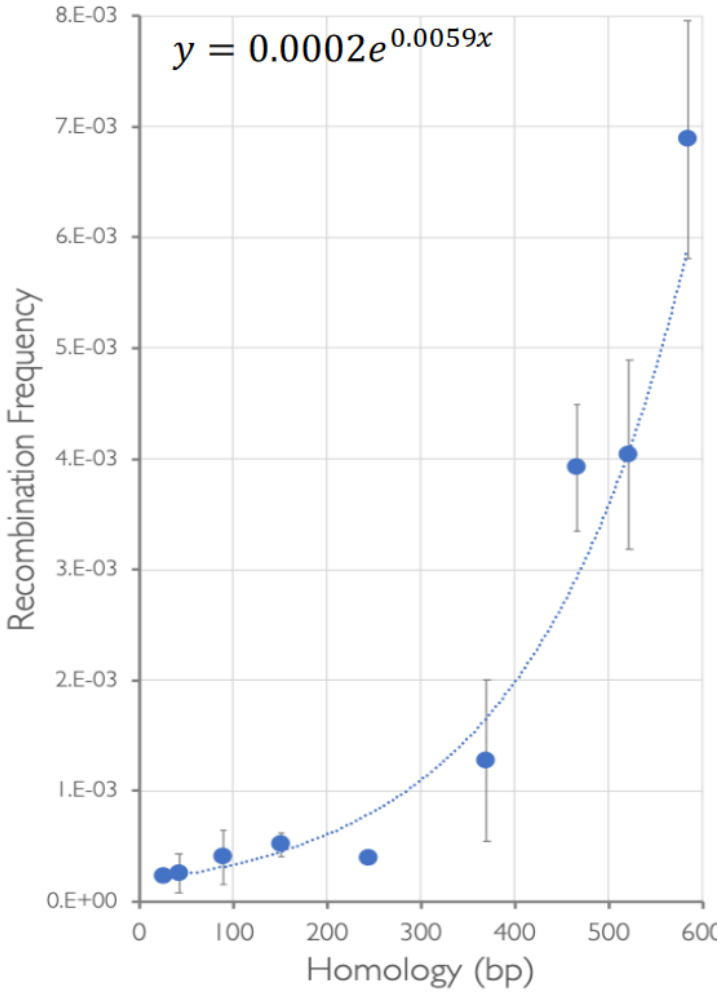


Figure 75. Frequency of recombination of truncated leuB *alleles in a wild-type* H. volcanii *background. From Jones (2019)*

While data pertaining to this relationship is not currently available for *H. mediterranei*, it may be assumed that a similar relationship exists in this species. In order to counteract the low recombination rate in *H. mediterranei*, deletion constructs with longer homology arms could be designed and used. A higher recombination rate during pop-out could result in a greater likelihood of isolating homoploid deletion mutants.

Alternatively, CRISPR could be used to attempt deletion of the gene of interest (as described in Stachler & Marchfelder, 2016). However, the intended alteration will still need to be confirmed to be present in all genome copies, so it is not guaranteed that this will improve the outcomes, depending on the essentiality of *hel308* in this species.

5.6.3 Determine essentiality through inducible expression

Smith (2021) proposed a novel scheme for introduction of an inducible promoter to chromosomal genes without deleting or disrupting the gene. This is potentially of great utility in the cases of essential genes or genes whose deletion incurs significant growth defects. The technique proposed is a variant of the pop-in/pop-out system, as shown in Figure 76, below.

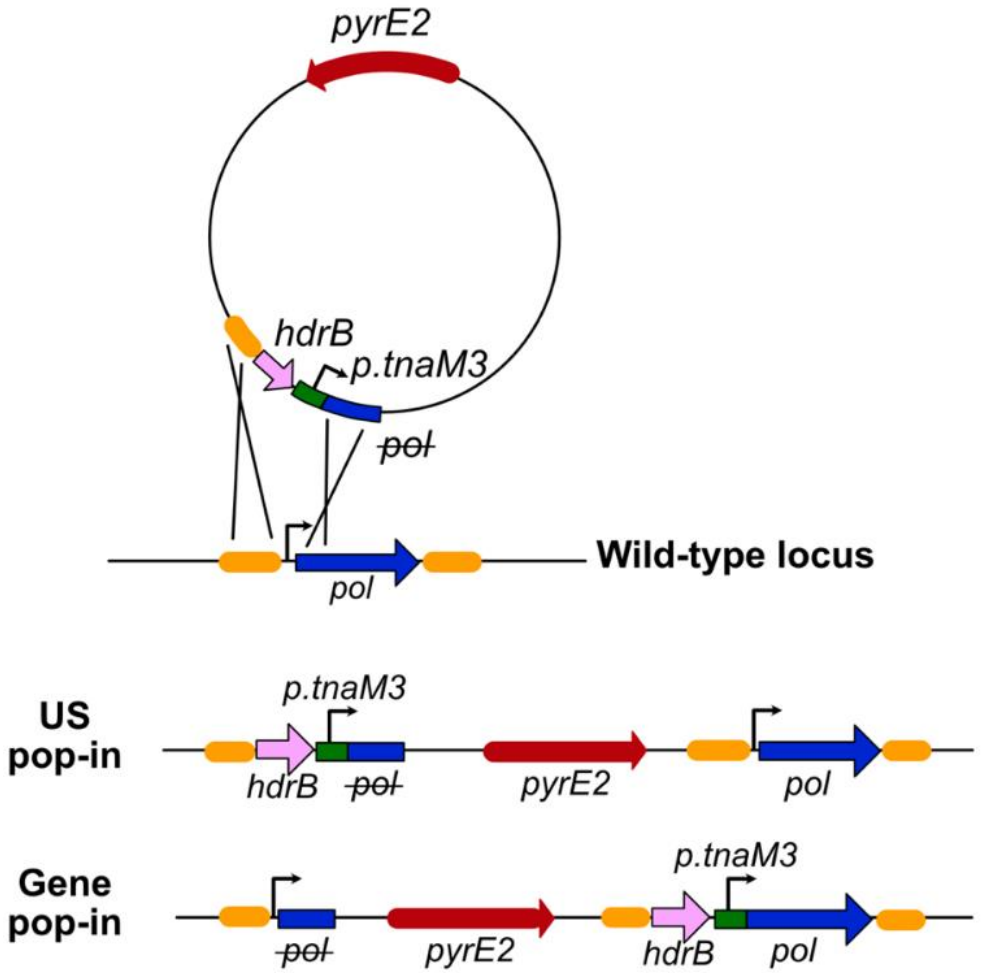


Figure 76. Strategy for placing chromosomal genes (in this case, "pol") *under the control of the inducible promoter* p.tnaM3. "*US" denotes upstream pop-in*. From Smith (2021)

Instead of placing the homology arms to either side of the gene of interest, the two homology arms are in this case an up-stream region and a truncated portion of the gene of interest. In the event of a pop-in in the gene, the gene is immediately placed under the control of the inducible promoter. Orientation of the pop-in can be checked as normal, and only pop-ins in the gene advanced to the next steps.

While phenotyping of the inducible gene of interest could potentially be carried out at this stage, there is a chance of the plasmid popping out and restoring the wild-type locus during growth. Although this could be combatted by maintaining selective pressure using the *pyrE2* and *hdrB* markers, pop-out of the vector is still recommended. Pop-out of the vector can be determined by resistance to 5'FOA due to loss of the *pyrE2* marker, as usual. However, the *hdrB* marker gene will remain present in colonies that have undergone an upstream pop-out, leaving the gene of interest downstream of both the *hdrB* marker and the inducible promoter, as shown in Figure 77, below.

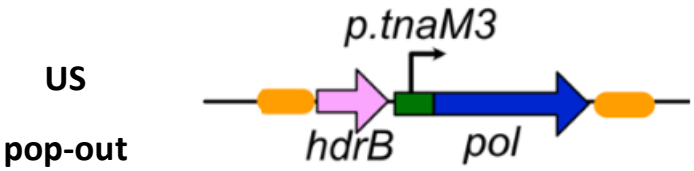


*Figure 77. Upstream ("US") pop-out event following gene pop-in results in the gene of interest ("*pol*") downstream of the* hdrB *marker and inducible promoter.*

As with a typical pop-in/pop-out scheme, presence of a thy+ phenotype in the pop-out strain is not necessarily indicative of homoploidy. Candidates would still have to be genotyped to avoid reversion to wild-type phenotype. This would ideally be performed through Southern blot. Should homoploid colonies be identified, the effects of gene expression and repression could be explored through the use of the inducible promoter (Large et al., 2007). If the gene under inducible control is in fact essential, cells should die in the absence of its expression, as seen in originless cells deprived of RadA (Hawkins et al., 2013).

However, this technique might be complicated somewhat in loci involving overlapping genes, as is the case with *hel308* and *cgi*. As these genes are likely transcribed together, placing an inducible promoter at the start of this operon could alter the expression of both genes. In this example, it would be preferable to instead delete the chromosomal copy while complementing with an inducible episome. Expression of the episomal copy would have to be maintained during the pop-out steps to combat any growth defect incurred by deletion of the chromosomal copy. Following confirmation of deletion, the effects of expression or repression of the episomal copy of the gene can be explored. Hopefully this strategy could elucidate the essentiality of *H. mediterranei hel308*.

5.7 Conclusion

H. mediterranei hel308 was previously claimed to be an essential gene, as attempts to delete it had consistently failed. Attempts to delete this gene under complementation in this study also produced no pop-outs, even where the gene was complemented with the wild-type allele, which should produce no growth defect. Further exploration revealed that the deletion vector used in the original reporting was flawed, and called into question the reported effect.

Re-design of the deletion vector also did not yield pop-outs in this study. However, the essentiality of *H. mediterranei hel308* has not yet been conclusively proven. Attempts to delete the *mrr* gene in this species also did not meet with success.

6. Interactions between origins, Hel308 and aphidicolin

6.1 Background

6.1.1 Replicative polymerases in the archaea

DNA polymerases are found in all organisms (as well as some viruses); these are organised into seven families; A, B, C, D, E, X and Y. The main replicative polymerases in eukarya belong to family B, while in bacteria this role is played by family C. Archaea generally contain two families of replicative DNA polymerases; family B, which is found in all archaea, and PolD, an enzyme which is unique to the archaea domain (but is absent from the crenarchaeal branch). This is summarised in Figure 78, below.

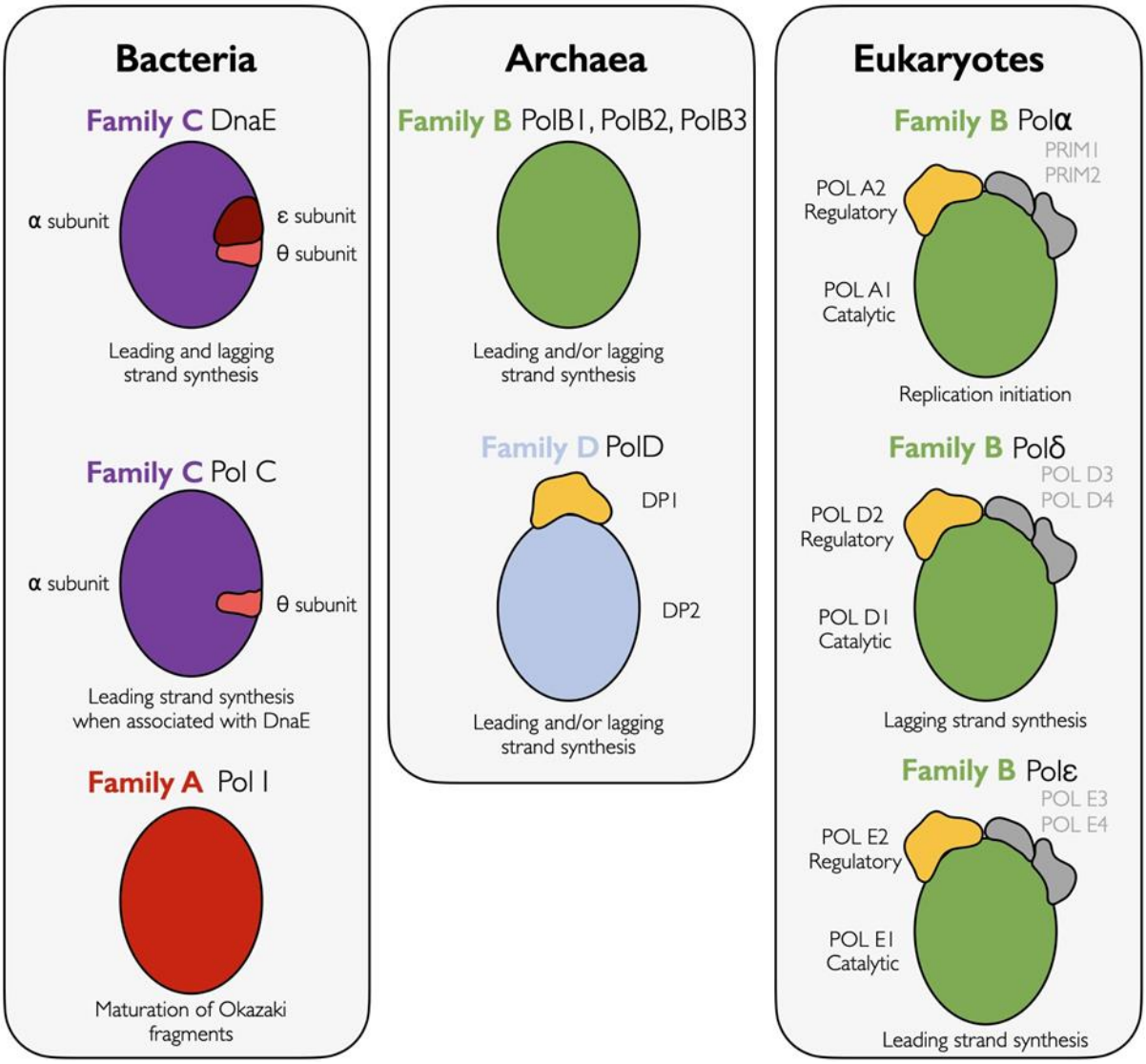


Figure 78. Representation of the major DNA polymerase families and their subunits found in the three domains. From T. Allers (personal communication).

6.1.1.1 Family B DNA polymerases

PolB DNA polymerases are widespread, being found in bacteria and Eukarya as well as archaea. In eukaryotes, multi-subunit family B DNA polymerases are responsible for genome replication; namely Pol α, Pol δ, and Pol ε. Pol α provides initial extension of an RNA primer, before "handing off" to the main replicative polymerases - Pol δ for the lagging strand, and Pol ε for the leading strand.

The core structures of all family B DNA polymerases are shaped like a "right hand", with domains including the "palm", "fingers" and "thumb". Archaeal DNA polymerases share this structure, and are considered ancestors of the eukaryotic replicative polymerases (Kazlauskas et al., 2020). In the archaea, these enzymes possess both DNA polymerase and exonuclease proofreading abilities, and are typically active as a monomer, rather than the multi-subunit affairs seen in the eukarya. The structure of an archaeal family B DNA polymerase complexed to a DNA strand is shown in Figure 79, below, demonstrating the palm, fingers, and thumb domains.

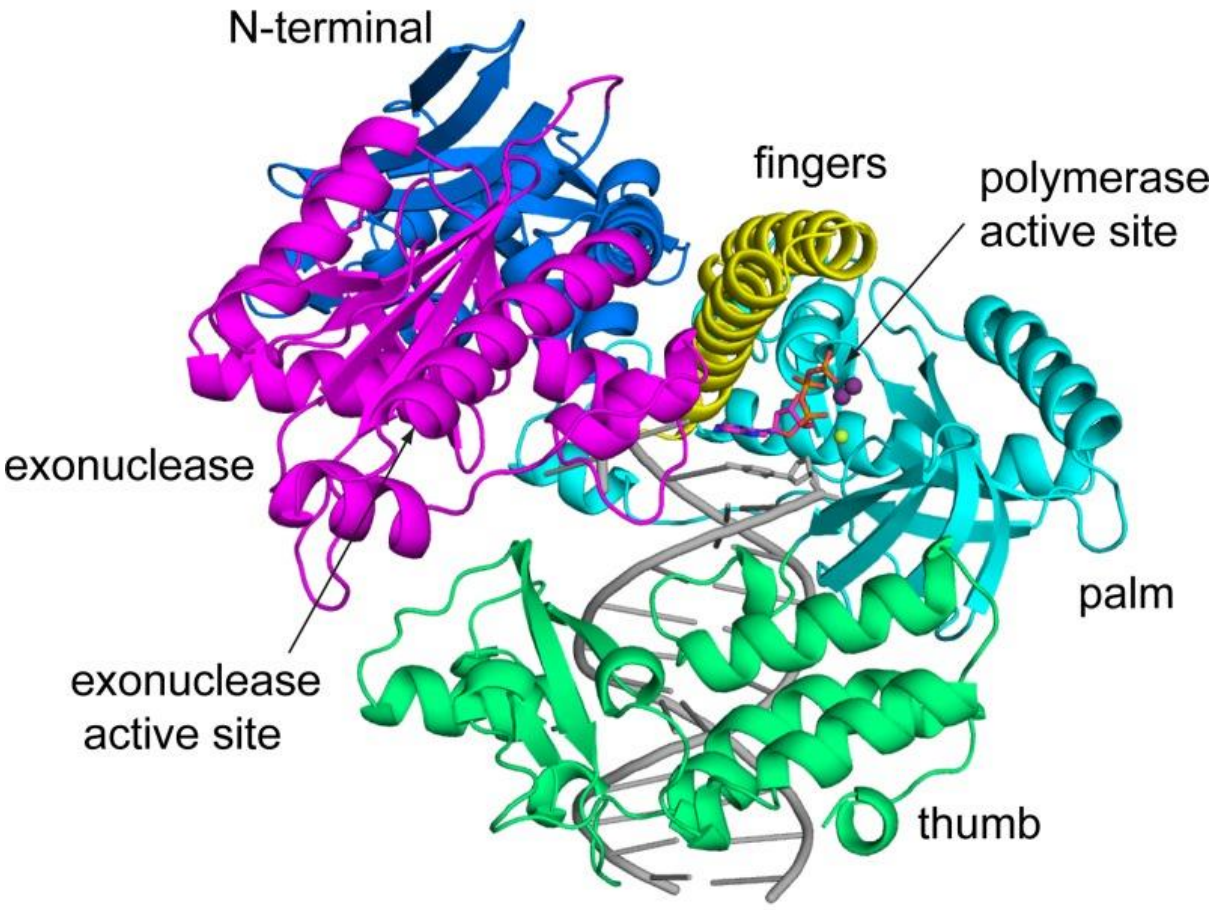


Figure 79. Structure of a family B DNA polymerase from Thermococcus kodakarensis. *The domains are shown in blue (N-terminal), magenta (exonuclease), yellow (finger), cyan (palm), and green (thumb). DNA is shown in grey, incoming dNTP as coloured sticks, and metal ions as spheres. Taken from Kropp et al. (2017).*

All known archaea species encode at least one copy of PolB, but multiple paralogues within the genome are common, particularly in the Crenarchaeota (Makarova et al., 2014). As Crenarchaeota possess no D family DNA polymerases, it is assumed that all of their DNA replication needs are able to be fulfilled by the various PolBs. Unlike family B DNA polymerases from bacteria and eukaryotes, archaeal PolB can detect deaminated bases, pausing 4 nucleotides before the damaged position (Killelea et al., 2010). This activity appears to be unique to members of this family found in the

archaea; it is not observed in bacteria (Wardle et al., 2008). In the archaea, PolB proteins are categorised into three families; PolB1, PolB2 and PolB3. The presence, dispensability, and inferred roles of these paralogues vary between clades.

PolB1 family DNA polymerases are found in all known crenarchaea, but are absent from the euryarchaeota (Makarova et al., 2014). *In-vitro* analysis of *Sulfolobus solfataricus* Dpo1 (a B1 family DNA polymerase) has been suggested to be the prime candidate for leading strand synthesis in this organism (Bauer et al., 2012). It is also transcribed at much higher levels than other DNA polymerases under normal growth conditions, and its activity is greatly bolstered *in vitro* by the addition of PCNA, RCF and SSB (Choi et al., 2011). However, it has been shown to struggle to synthesise past some kinds of DNA lesions, perhaps suggesting specialised roles for the other paralogues in this species.

Members of the PolB2 family are usually considered functionally inactive in many organisms, due to mutations to catalytic residues (Rogozin et al., 2008). However, Dpo2, a *Sulfolobus islandicus* DNA polymerase of the PolB2 family, has been shown to be a functional and high-fidelity DNA polymerase, despite lacking exonuclease activity. It effectively extends DNA synthesis from nucleotide mismatches opposite DNA lesions, but is inefficient at bypassing template lesions itself (Feng et al., 2022). This DNA polymerase activity is distributive, wherein the enzyme synthesises relatively few nucleotides before dissociating from the strand. This is suggestive of a role in an extension step of translesion synthesis and DNA damage tolerance. However, it is unknown whether these abilities are present in homologues outside the crenarchaeota.

Phylogenetic analysis of PolB3 homologues within the archaea reveals that their evolutionary relationships do not always match those of their host species (Makarova et al., 2014). In addition to this, some species carry multiple PolB3 homologues, seemingly from different sources. This suggests multiple instances of horizontal gene transfer within the domain, or possibly cases of accelerated evolution in some phyla.

In *Sulfolobus solfataricus,* the B3 family DNA polymerase Dpo3 has been suggested to play a role in lagging strand synthesis (Bauer et al., 2012) or DNA repair (Bohall & Bell, 2021). While not being essential, absence of this protein confers changes to replication profiles, and increased susceptibility to some kinds of DNA damage (Bohall & Bell, 2021).

In the euryarchaeon *Thermococcus kodakarensis,* the sole *polB* gene (belonging to the PolB3 family) is inessential, and can be deleted without negatively impacting growth rate (Kushida et al., 2019). This suggests that PolD, rather than PolB, is responsible for genome replication in this species. However, strains bearing the *polB* deletion showed increased sensitivity to a range of DNA damage agents, implying an important role in repair pathways in this organism.

Haloferax volcanii possesses two *polB* genes. These are *polB1*, belonging to the B3 family, and *polB2*, which is a B2 family DNA polymerase. *PolB2* is carried on the pHV4 minichromosome, and has been shown to be inessential (T. Allers, unpublished research). The high percentage of rare codons in this gene, and its absence from the closely related *H. mediterranei* suggest that it is a recent acquisition by horizontal gene transfer; possibly viral in origin (Smith, 2021). It is therefore not thought to play an integral role in genome replication.

The other *H. volcanii polB* gene, *polB1,* is carried on the main chromosome. This gene has been shown to be essential (Smith, 2021). However, it also carries a large intein; a protein component that removes itself from the finished protein by self-splicing. Inteins are found in PolB3 in many archaea, sometimes as many as three within the same gene (Naor et al., 2011). The single intein present in *H.*

volcanii PolB1 is 437 amino acids long, and can be deleted without significantly impacting the growth rate of the organism.

6.1.1.2 Family D DNA polymerases

PolD is an enzyme unique to the archaea, being present in all archaeal phyla save for the crenarchaeota; it is therefore suggested that this enzyme was present in the last archaeal common ancestor (LACA) (Makarova et al., 2014). Unusually, this DNA polymerase functions as a heterodimer, being made up of two subunit proteins; PolD1 and PolD2 (denoted PolD-S and PolD-L in some species) (Greenough et al., 2015). The small PolD1 possesses proofreading and 3'-5' exonuclease activities, while the larger PolD2 forms the catalytic core of the enzyme.

PolD1 belongs to a calcineurin-like phosphoesterase family of proteins, and exhibits greatest similarity to Mre11, an exo/endonuclease involved in DNA repair (Sauguet et al., 2016). It is composed of two domains; the exonuclease domain, and the ssDNA-binding OB-fold. It is thought to be an ancestor of one of the eukaryotic family B DNA polymerase subunits, which have since lost their exonuclease activity (Gueguen et al., 2001). As a 3'-5' exonuclease, PolD1 demonstrates a preference for 5' overhangs, and is more active where mismatches are present, thus removing the offending bases. This is consistent with its role as a proofreading subunit (Jokela et al., 2004). Interestingly, it is fully capable of this activity *in vitro*, even without the DP2 subunit present.

X-ray crystallography has revealed that the core of PolD2 contains a double-psi β-barrel, a structure associated with the "two-barrel" family of RNA polymerases (RNAPs) (Sauguet et al., 2016). "Two-barrel" RNAPs are found in all domains of life, as well as a few viruses, and have been observed acting as transcriptases as well as in microRNA amplification. It is an interesting point that these two groups of enzymes with different functionalities may share a common ancestor.

Haloferax volcanii possesses a single PolD homologue, encoded by the genes *polD1* and *polD2. polD1* is located very close to the *oriC1* origin, while *polD2* is located further away. *polD2* has been shown to be essential (Smith, 2021).

6.1.1.3 Roles of B and D polymerases

The variation in the DNA polymerase enzymes present in different clades of the archaea can make it difficult to predict the roles that each will play in a given species. In *S. solfataricus*, the B2 and B3 family DNA polymerases are both dispensable, although this may confer susceptibility to certain kinds of DNA damage (Bohall & Bell, 2021). B1 is thus the dominant replicative polymerase in this species, which contains no PolD. Meanwhile in *Methanococcus maripaludis*, which carries only PolD and PolB3 DNA polymerases, PolD has been signposted to be essential, while PolB is thought to be inessential (Sarmiento et al., 2013). The same has been confirmed to be the case in *Thermococcus kodakarensis* (Cubonová et al., 2013).

In a minimal Okazakisome model in *Thermococcus* species 9°N, PolB has been shown to exhibit very weak activity when extending RNA primers, appearing to prefer extending from a DNA primer (Greenough et al., 2015). PolD has been suggested to be the main replicative polymerase in this species; however, it appears to be inhibited by the presence of downstream Okazaki fragments. PolB was observed filling the gaps left by PolD synthesis, displacing the RNA primer to form a flap structure. However, it should be considered that this experiment was performed using a "minimal Okazakisome" consisting of PolB, PolD, PCNA, Fen1, and DNA ligase in various combinations. It is possible that additional factors act *in vivo* to modify the activity of these DNA polymerases.

In *Pyrococcus abyssi*, PolD is thought to be involved in the initial extension of RNA primers, before "handing off" to PolB, which is capable of displacing it from the DNA strand (Rouillon et al., 2007). As

PolB is inhibited by collision with downstream primers, this may be a phenomenon more common to the leading strand; or could implicate PolD in maturation of Okazaki fragments (Henneke et al., 2005). More recent research in this species has shown that both PolB and PolD are capable of extending RadA-mediated recombination structures, although PolB performs this role with greater efficiency (Hogrel et al., 2020). This is consistent with a suggested role for PolB in DNA repair in this organism, but also explains its inessentiality, as PolD is also capable of this function. As before, it is possible that additional replisome components could increase the effectiveness of this process *in vivo*.

In *Pyrococcus furiosus,* both PolD and PolB contain C-terminal PIP-boxes, allowing their interaction with PCNA. The presence of PCNA (and its loader, RFC) increases DNA polymerase activity in both enzymes significantly (Tori et al., 2007). The *H. volcanii* PolD is highly similar to its *P. furiosus* homologue, and is likely also capable of this interaction (MacNeill, 2009). A summary of the presence and absence of different polymerase families in difference archaea is presented in Figure 80, below.

While *H. volcanii*'s *polB2* has been shown to be dispensable, *polB1* and *polD* are both known to be essential. This is a similar pattern to *Halobacterium*, in which *polD* and *polB1* are essential, but the minichromosomally-encoded *polB2* is inessential (Berquist et al., 2007). However, the exact roles of these two essential enzymes are not yet elucidated in *H. volcanii*.

		PolD	PolB1	PolB2	PolB3
Crenarchaea	*Sulfolobus solfataricus*	○	●	●	●
	Sulfolobus islandicus	○	●	●	●
Euryarchaea	*Haloferax volcanii*	●	○	●	●
	Haloferax mediterranei	●	○	○	●
	Halobacterium salinarum	●	○	●●●	●
	Thermococcus sp.	●	○	○	●
	Pyrococcus sp.	●	○	○	●
	Methanococcus maripaludis	●	○	○	●

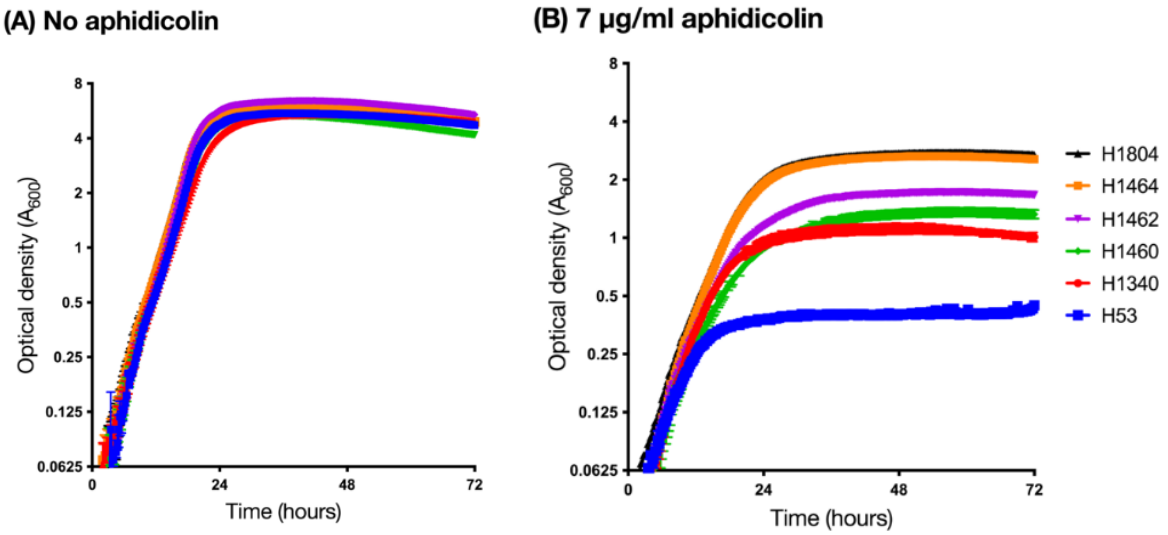
Figure 80. Summary of the presence of different DNA polymerases in the species discussed in this chapter. ○: not present. ●: present (indicates numbers).

6.1.2 Aphidicolin and origins in *Haloferax*

Aphidicolin is an antimicrobial drug that can inhibit the DNA polymerase activity of family B DNA polymerases, but not their exonuclease activity. It is a tetracyclic diterpenoid, isolated from the mould *Cephalosporium aphidicola*. X-ray crystallography of aphidicolin inhibiting human Pol α has revealed this drug's method of effect (Baranovskiy et al., 2014). Aphidicolin enters and occupies the active site within the enzyme which would usually be occupied by the base and sugar (but not the phosphate groups) of an incoming dNTP during DNA synthesis. This prevents entry by a relevant dNTP, and also mispositions the opposite base in the template strand. This interaction is thought to be transient; however, as the x-ray crystallography shows the enzyme still interacting with both the DNA and aphidicolin, it seems that this may result in a stalled DNA polymerase, thus arresting further DNA synthesis.

Smith (2021) reported that deletion of origins positively correlates with increased resistance to treatment with aphidicolin. As aphidicolin specifically inactivates PolB family DNA polymerases, this suggests a reduced usage of PolB in these strains; adding further weight to the idea that PolB is the

dominant DNA polymerase in origin-dependent replication, while PolD play a central role in origin-independent replication. These data are shown in Figure 81, below.



Strain	Genotype
H53	*ΔpyrE2, ΔtrpA,*
H1340	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2*
H1460	*ΔpyrE2, ΔtrpA, ΔoriC2, ΔoriC3*
H1463	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC3*
H1464	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*
H1804	*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3, Δori-pHV4-2*

Figure 81. Impact of aphidicolin on Haloferax *strains. The greater the number of origins deleted, the greater the resistance to aphidicolin treatment. From Smith (2021).*

This is in line with data from *Thermococcus kodakarensis*. While this species contains both PolB and PolD, PolB is inessential, and its deletion or overexpression does not affect the growth rate of the strain (Kushida et al., 2019). PolD, meanwhile, is essential (Cubonová et al., 2013). This likely indicates that PolD is the main replicative polymerase in this species. Furthermore, deletion of the sole *orc1/cdc6* homologue in *T. kodakarensis* has demonstrated that it is inessential, and replication profiles generated by marker frequency analysis show that no peaks corresponding to unidentified or dormant origins are found in the resulting strain (Gehring et al., 2017). However, examination of replication profiles of the wild-type *T. kodakarensis* shows an identical profile, meaning that origins are not used for replication even when present. The inability to delete either *radA* or *radB* in this species suggests that recombination-based methods are responsible for genome replication even when origins are present (Gehring et al., 2017). Together, these data suggest a broader pattern wherein PolB is the main DNA polymerase involved in origin-dependent genome replication, while PolD is involved in recombination-based replication, hence the dispensability of both *polB* and *cdc6* in this species.

However, it appears that PolB must still be playing some role within origin-less *H. volcanii*, as the *polB1* gene remains essential even in a *Δori* background (Smith, 2021) (although the strain used in this experiment still contained the pHV1 and pHV3 minichromosomes, which still possessed their origins). Deletion of *polB* from *T. kodakarensis* has revealed that it plays an important role in DNA

repair (Kushida et al., 2019). It is possible that secondary roles may yet be identified for this enzyme in *H. volcanii.*

In *H. volcanii*, preliminary data has shown that deletion of *hel308* causes increased sensitivity to aphidicolin (Dattani, personal communication). The full interactome of Hel308 is not yet fully characterised, so the exact cause of this effect is not yet known. Pulldown experiments using tagged Hel308 yielded a range of replisome components, but PolD1 and PolD2 were not among them (Lever, 2020).

6.2 Aims and objectives

Deletion of origins causes aphidicolin resistance, while deletion of *hel308* causes aphidicolin sensitivity. It was not known which of the two effects would be dominant in *Δhel308 Δori* cells. This seemed a good opportunity to explore the nature of the interactions between Hel308, origin deletion, and aphidicolin resistance.

In addition to the deletion of *hel308*, several *hel308* point mutants were also selected, to see if this would shed light on any qualities necessary for the interaction or role that Hel308 plays. Several point mutations were selected for investigation; ATPase-null RecA domain mutants K53A and D145N, and hyper-helicase point mutant F316A. These would inform whether the helicase activity of Hel308 is important for this interaction, or whether Hel308 helps to target or recruit other factors in the cell.

6.3 Aphidicolin response in origin-deleted strains

6.3.1 Strains construction

Aphidicolin assays were undertaken in strains with intact origins, as well as strains lacking chromosomal origins. The strain lacking all chromosomal origins was H1804 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*). However, replacing the chromosomal *hel308* copy with the point mutants would be difficult in this strain, as it would be challenging to ascertain that the wild-type gene had been replaced in all copies. Under these circumstances, it is more convenient to delete the gene of interest entirely, and then re-introduce a point mutant allele. This allows surety of the presence of the point mutant in all genome copies, which would otherwise be difficult to determine.

Chromosomal replacement vectors for each of the three point mutants of interest already existed; p1335 for *hel308-D145N* (produced by T. Allers. *dam*- preparation p1337), p1642 for *hel308-F316A* (produced by R. Gamble-Milner. d*am*- preparation p1647), and p1988 for *hel308-K53A* (produced by B. Lever. *dam*- preparation p2015).

A strain in which the *hel308* gene had been deleted already existed; H5366. As the *hel308* deletion was marked with the *trpA* gene, replacement of this locus with the various point mutants could be detected via tryptophan auxotrophy. This would indicate that all chromosomal copies bear the desired allele, and that the strain cannot revert to a different phenotype during the experiment. For added security, the homoploidy of each strain was confirmed by Southern blot.

For each candidate tested by Southern blot, gDNA was digested with the NcoI and MluI restriction enzymes overnight. The positions of these sites in the wild type, *Δhel308::trpA* and *Δhel308* alleles of this locus are shown in Figure 82, below.

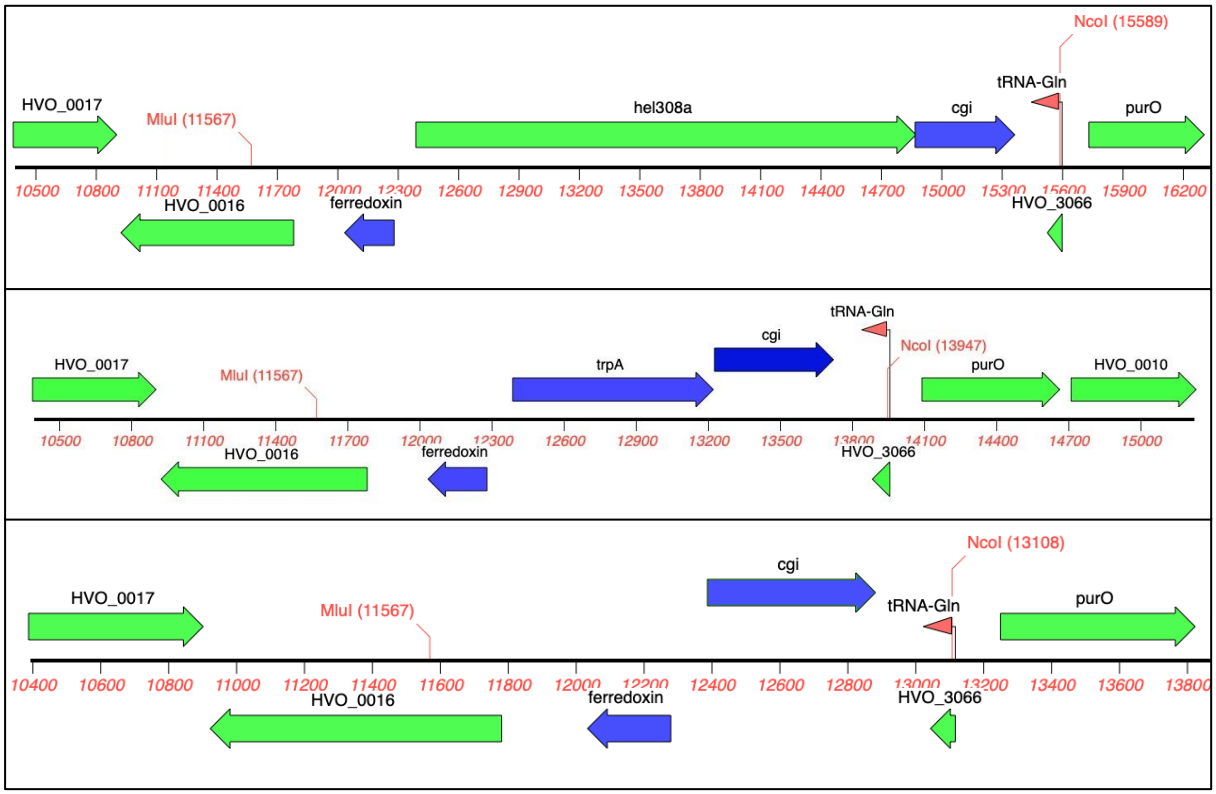


Figure 82. Position of the NcoI and MluI restriction sites at the hel308 *locus in wild-type (top),* Δhel308::trpA *(middle) and* Δhel308 *(bottom) alleles.*

The expected band sizes produced by this digestion are expected to be 4kb (wild-type), 2.4kb (*trpA*-marked *hel308* deletion) and 1.5kb (unmarked *Δhel308*). The Southern blots confirming the presence of the *hel308* gene for each point mutant are shown in Figures 83 and 84, below.

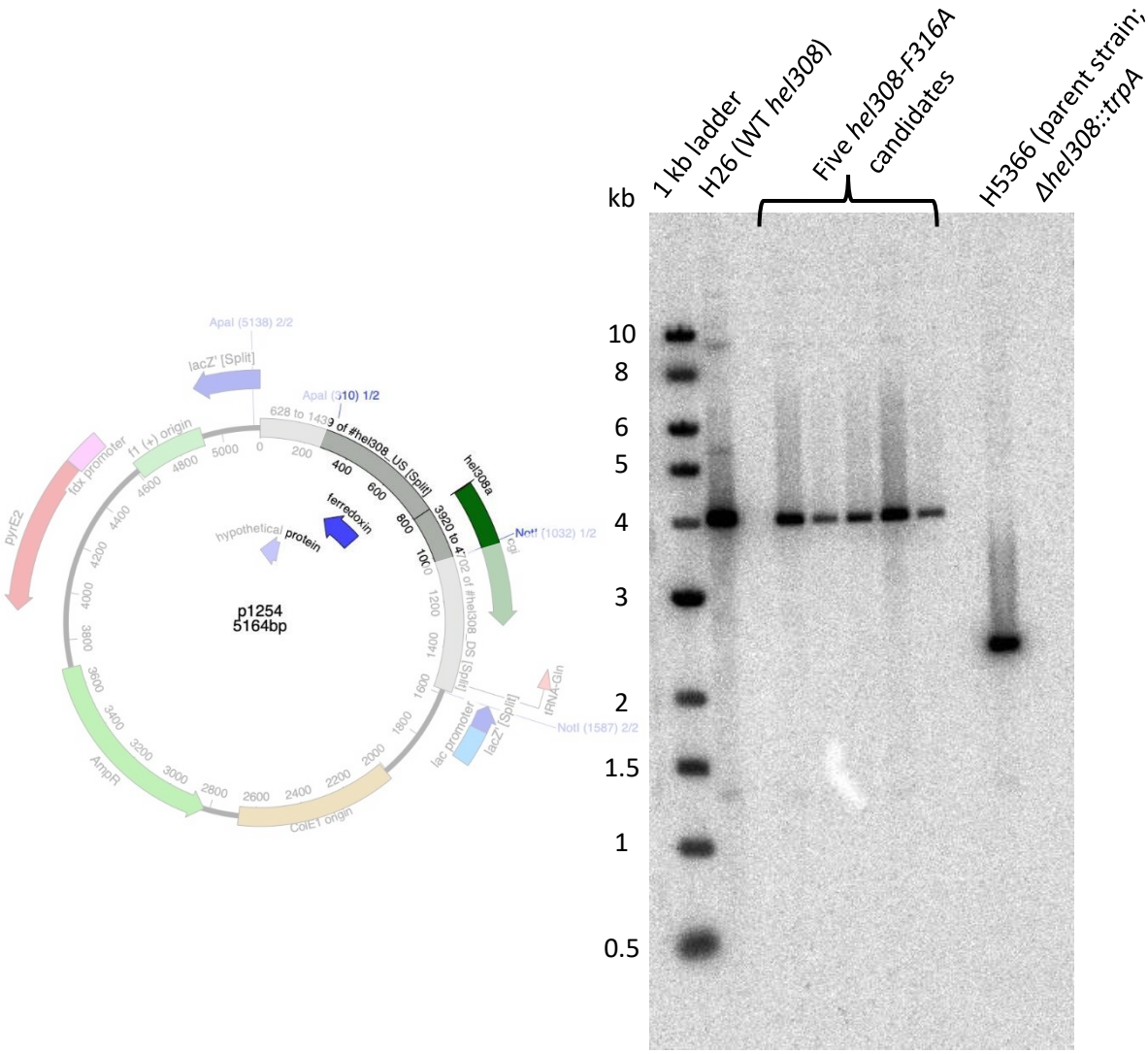


Figure 83. Southern blot confirming the addition of the hel308-F316A *allele to the H5366 background. Genomic DNA had been digested with MluI and NcoI. Probe was produced by digestion of p1254 with ApaI and NotI, with the 700bp fragment being excised and purified.*

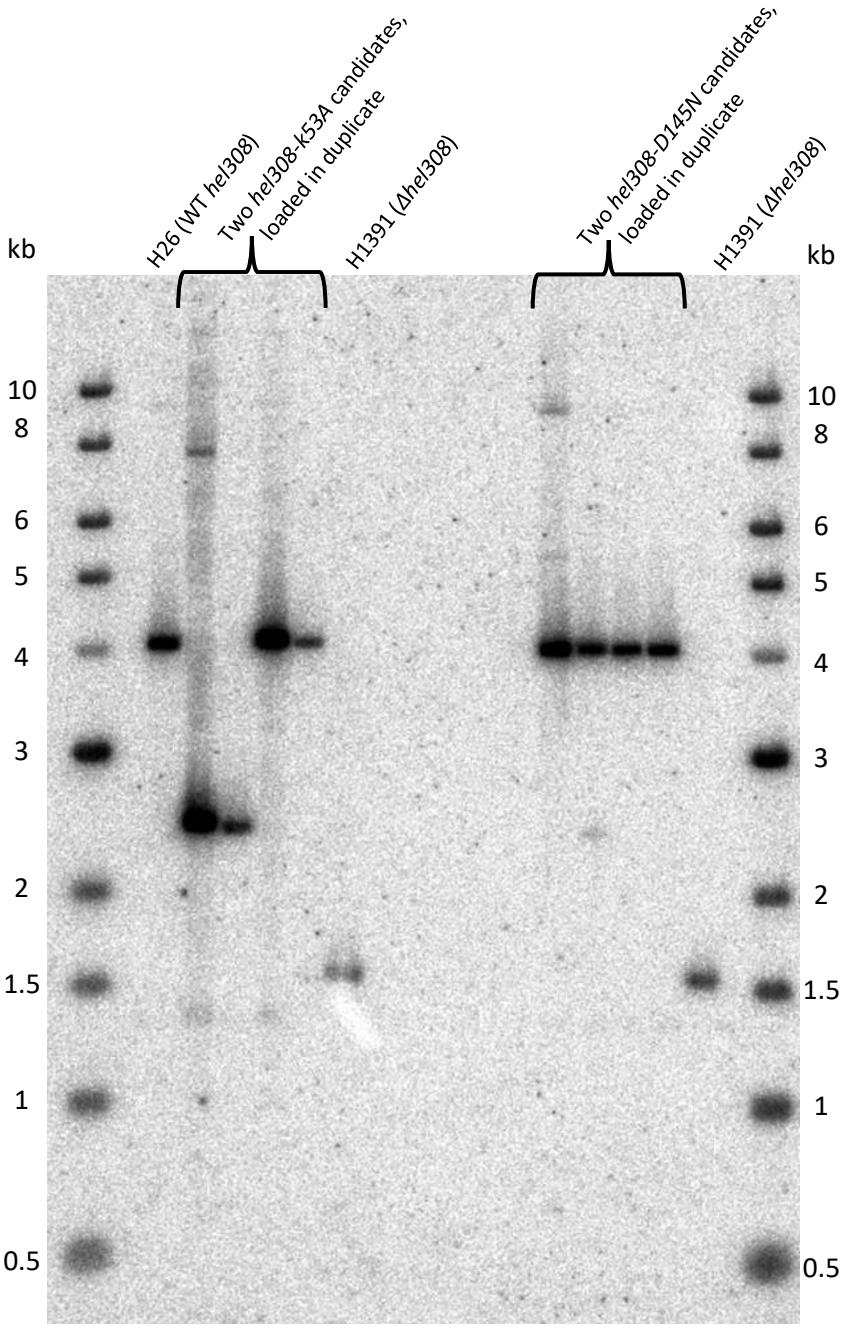


Figure 84. Southern blot confirming the addition of the hel308-K53A *and* hel308-D145N *alleles to the H5366 background. Genomic DNA had been digested with MluI and NcoI. Probe used was the same as that described above.*

In the above Southern, it is apparent that one of the two *hel308-D145N* candidates was in fact merodiploid, producing bands at both 4kb (corresponding to full-length *hel308*) and 2.4kb (corresponding to *Δhel308::trp*). This emphasises the importance of confirming genotype by Southern blot, and not relying on phenotyping to confirm genetic alterations.

A *Δhel308* control was also desired for comparison with the point mutant strains, but H5366 was unsuitable for this purpose, as the presence of the *trpA* gene could confer a small growth benefit compared to the other strains (in which the *trpA* marker had been removed during addition of the point mutants). This marker was therefore deleted from H5366 using the p1276 deletion vector, shown in Figure 85, below.

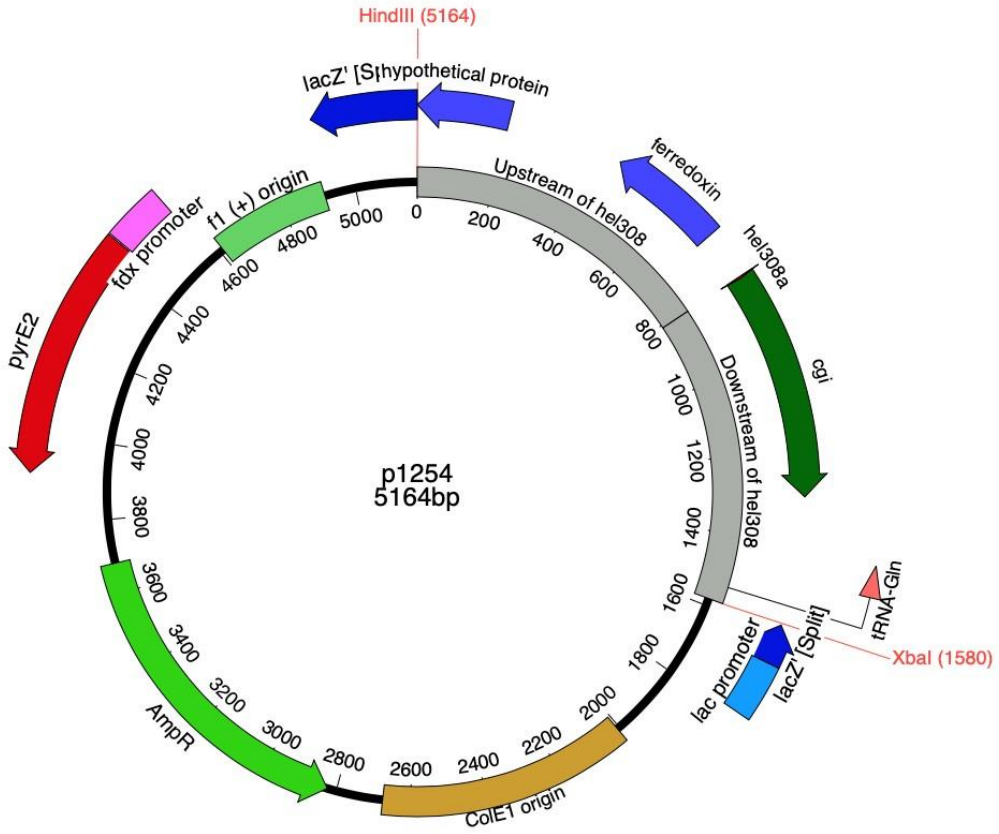


Figure 85. The hel308 *deletion vector p1254, produced by Thorsten Allers (p1276 is the designation of the dam⁻ preparation of this plasmid).*

As with the introduction of the *hel308* point mutants into H5366, removal of the *trpA* marker could be detected by tryptophan auxotrophy. Deletion of the *trpA* gene from the locus was also confirmed by Southern blot, to ensure homoploidy. This is shown in Figure 86, below.

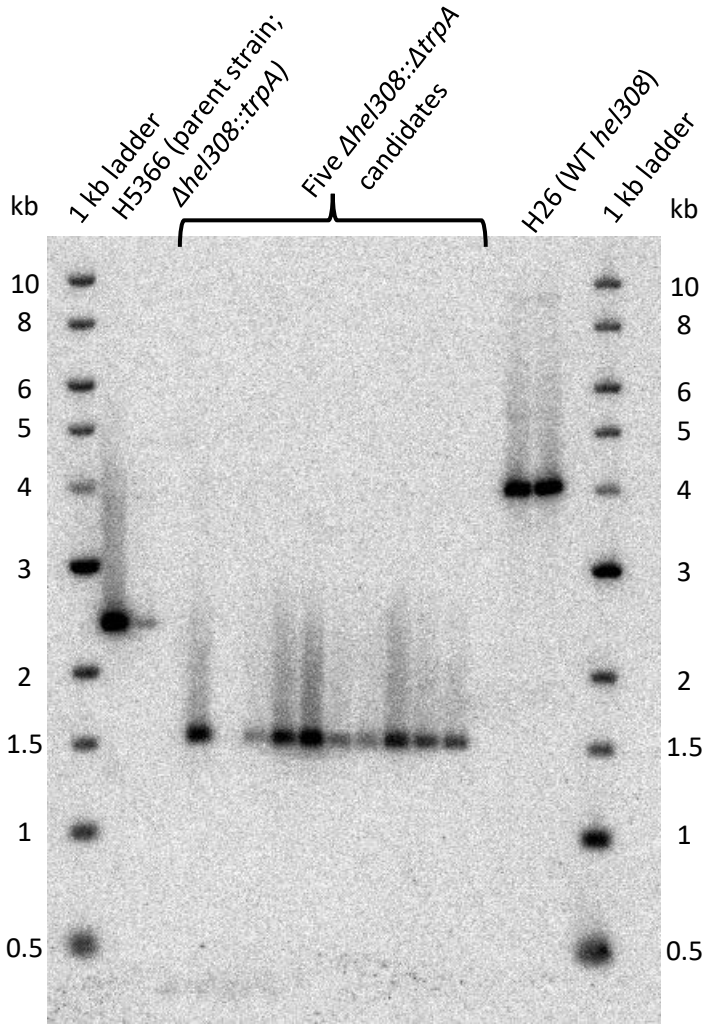


Figure 86 Southern confirming deletion of trpA *marker from* hel308 *locus of H5366. Genomic DNA had been digested with MluI and NcoI, with digestions performed and loaded in duplicate. Probe used was the same as that described above.*

A summary of the strains used in the aphidicolin assays are shown in the table below:

Table 6-1 Strains used in this section

Strain	**Genotype**	**Notes**
H1804	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4*	Produced by Katarzyna Ptasinska. Ancestor of the strains listed below; used in the aphidicolin assay as a wild-type *hel308* control.
H5366	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4*	Produced by Ambika Dattani. H1804 with *hel308* deleted, leaving a *trpA* marker in its locus.

	Δhel308::trpA	
H5736	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4* *Δhel308*	Produced for this study. H5633 with the *trpA* marker removed using the p1276 vector.
H5725	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4* *hel308-K53A*	Produced for this study. H5366 with the *hel308-K53A* allele introduced via the p2015 vector.
H5720	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4* *hel308-D145N*	Produced for this study. H5366 with the *hel308-D145N* allele introduced via the p1335 vector.
H5714	*ΔpyrE2,* *ΔtrpA,* *ΔoriC1,* *ΔoriC2,* *ΔoriC3* *Δori-pHV4* *hel308-F316A*	Produced for this study. H5366 with the *hel308-F316A* allele introduced via the p1647 vector.

6.3.2 Response to aphidicolin in *Δori* strains

A growth assay was carried out in the presence of aphidicolin. As chromosomal origins are deleted in the strains used, strong aphidicolin resistance was expected. Therefore 7µg/ml and 12.5µg/ml of aphidicolin (or an equivalent amount of DMSO as a control) were used, which allows a good range of responses to be observed in *Δori* strains (L Mitchell, personal communication). While *H. volcanii* can metabolically interact with DMSO, the genes that make this possible are only transcribed under anaerobic conditions, allowing growth by DMSO respiration (Qi et al., 2016). The presence of DMSO should therefore not prove a confounding factor.

As three conditions were examined for each of five strains, the results are difficult to distinguish on a single graph. They are therefore shown in Figures 87 and 88 below, separated first by strain and then by aphidicolin concentration. Data shown is the result of three biological replicates per strain, per condition.

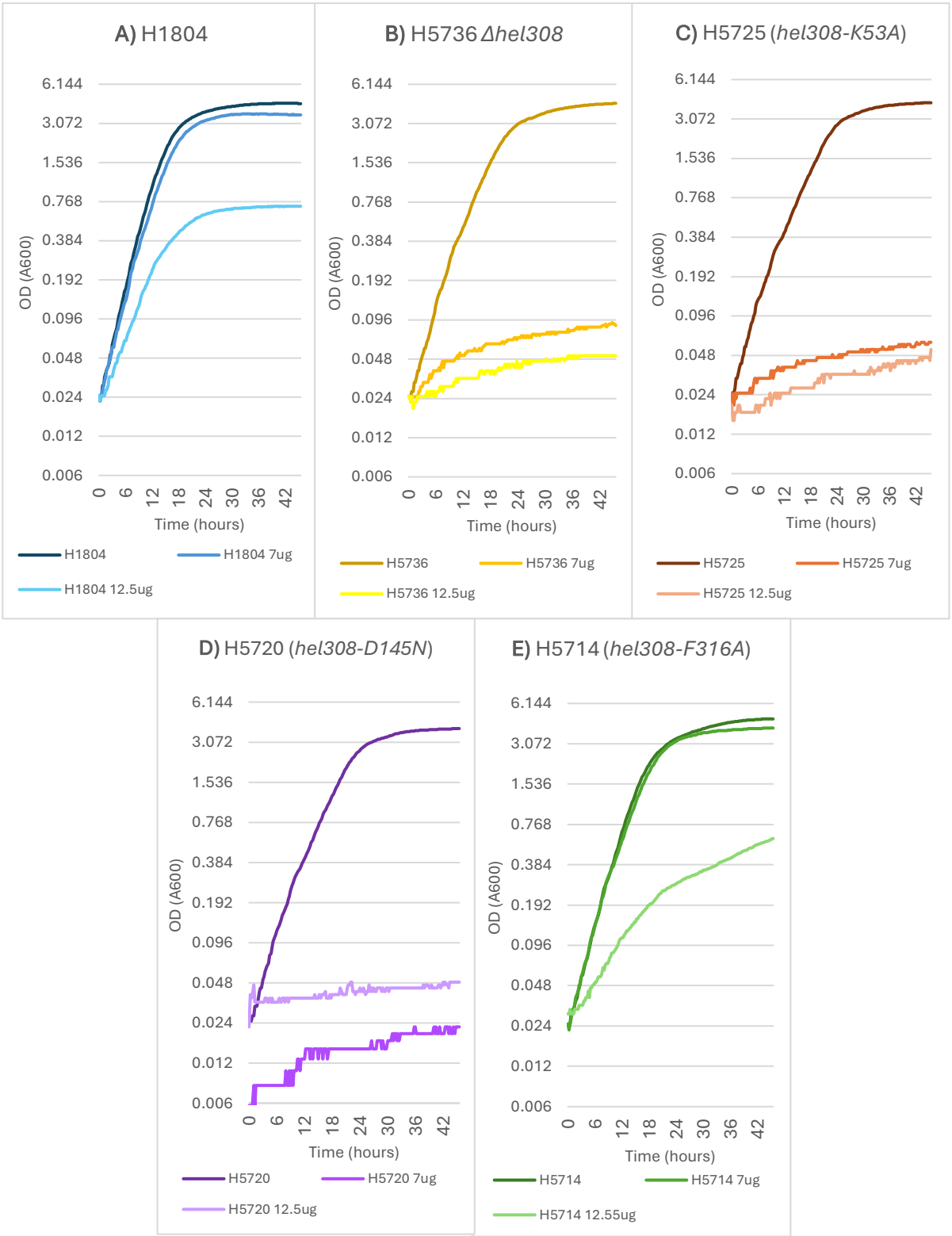


Figure 87. Effect of aphidicolin on growth of originless strains, separated by strain.

Graphs comparing growth rates of the five different strains under different concentrations of aphidicolin are shown below:

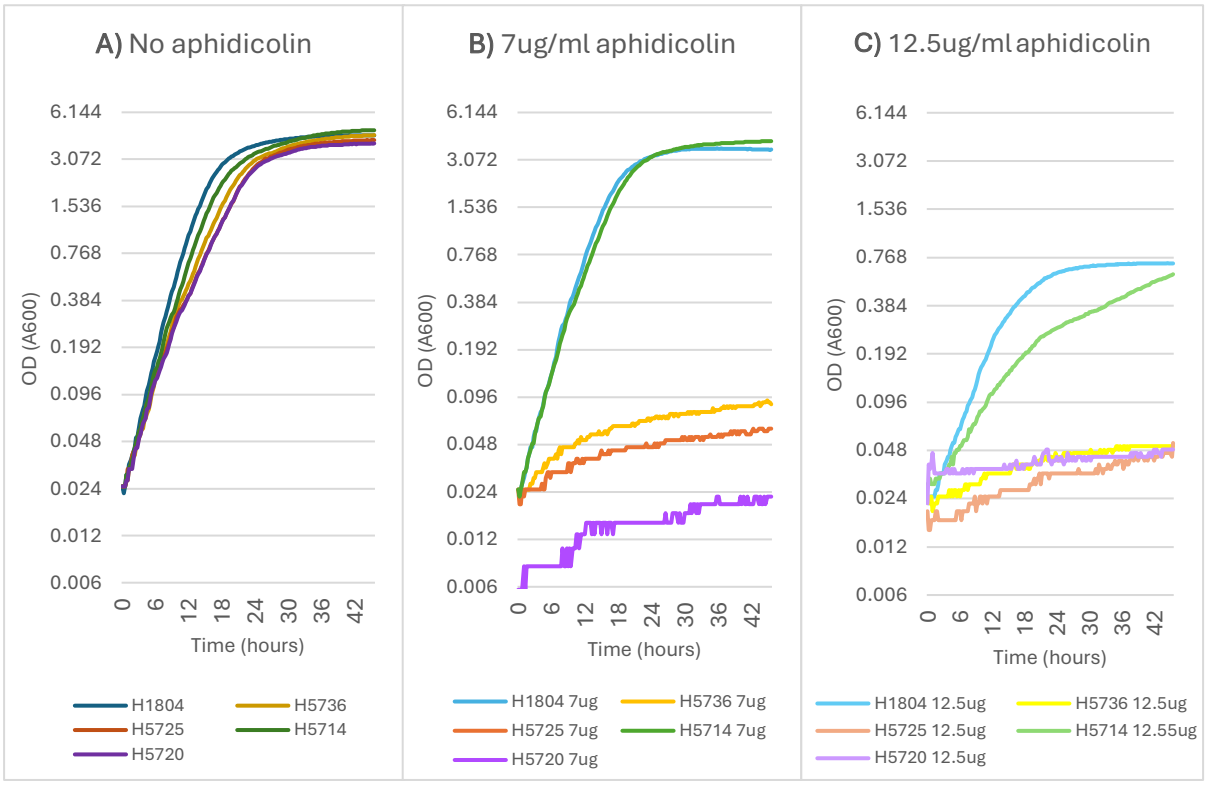


Figure 88. Effects of aphidicolin concentration on the growth of originless strains, separated by aphidicolin concentration.

Some anomalies are seen in the above data with regards to strain H5720, specifically in figure 87D and Figure 88B. This is an artefact of how the data are processed for this assay. In order to ensure that the strains are synchronised at the start of log phase at the start of the graph, an arbitrary cutoff is selected; in this case the first value for each well to report an OD above 0.02. All data points prior to this are discarded, as the signal-to-noise ratio is generally very low before this point. However, in strains most severely affected by the aphidicolin, ODs had still not reached the cutoff value at the end of the 72-hour assay. The data from these wells are still included, if only to demonstrate their extremely slow growth rate.

As shown in Figure 88A, growth rate in the absence of aphidicolin is broadly similar, and all five strains enter stationary phase at an OD of between 3.8 and 4.3. H5714 enters stationary phase at a slightly denser OD; around 4.7, suggesting that hyper-helicase Hel308 may be confer a small benefit to the cell.

As can be seen in Figure 88, two phenotypes are evident across the five strains; a Hel308-wild-type-like, which demonstrates strong aphidicolin resistance, and a *Δhel308*-like phenotype, which does not tolerate aphidicolin well. The ATPase-null Hel308 point mutants behave similarly to *Δhel308* strains in the context of aphidicolin, suggesting that resistance to this chemical involves the ATPase or helicase activity of this enzyme. As expected, the hyper-helicase point mutant Hel308-F316A (H5714) behaves similarly to the wild-type, and may even exhibit a small growth advantage over H1804 (wild-type Hel308) in the no-aphidicolin and 7µg/ml aphidicolin. It is also noticeable that, at the end-point of the 12.5µg/ml assay, H5714 was still growing, and appeared to be in (slow) exponential phase, while H1804 had reached a stationary OD. Had the assay continued, it is possible that the Hel308-F316A mutant may have eventually reached a much higher stationary phase.

While H1804 and H5714 are not severely impacted by the 7µg/ml aphidicolin treatment, they do clearly struggle to grow under the 12.5µg/ml aphidicolin, suggesting some requirement for PolB activity remains in these strains. Alternatively, this could be due to toxicity or off-target effects of the aphidicolin. However, it should be considered that, while all four chromosomal origins have been deleted in this background, the megaplasmids pHV1 and pHV3 are still present, and their origins are still intact. Is it possible that the growth defect could therefore be due to difficulty in replicating pHV3, which is essential and (for reasons yet unknown) cannot be replicated by origin-independent mechanisms (Marriott, 2018).

As shown in Figure 88A, growth in YPC is broadly similar for all five of these strains. The marked difference between the strains under treatment with aphidicolin seems to suggest that Hel308 would usually play an important role in origin-independent replication in the absence of PolB, and that disruption of this pathway adversely affects growth to a severe degree. This is potentially indicative of a shared pathway including Hel308 and PolD. Pulldown experiments have not previously demonstrated a direct interaction between these two proteins (Lever, 2020), but this does not rule out an indirect interaction or association between the two.

6.3.3 Response to low-dose aphidicolin

The growth inhibition in *Δhel308*-like strains was greater than expected, resulting in extremely low growth in both of the aphidicolin treatments. The assay was therefore repeated with lower aphidicolin concentrations (2.5 and 5µg/ml) to see whether this allowed any differences to be observed between the *Δhel308* and ATPase-null mutants. As before, the data are shown in Figures 89 and 90 below, grouped first by strain, and then by growth condition. Three biological replicates were included per strain, per condition.

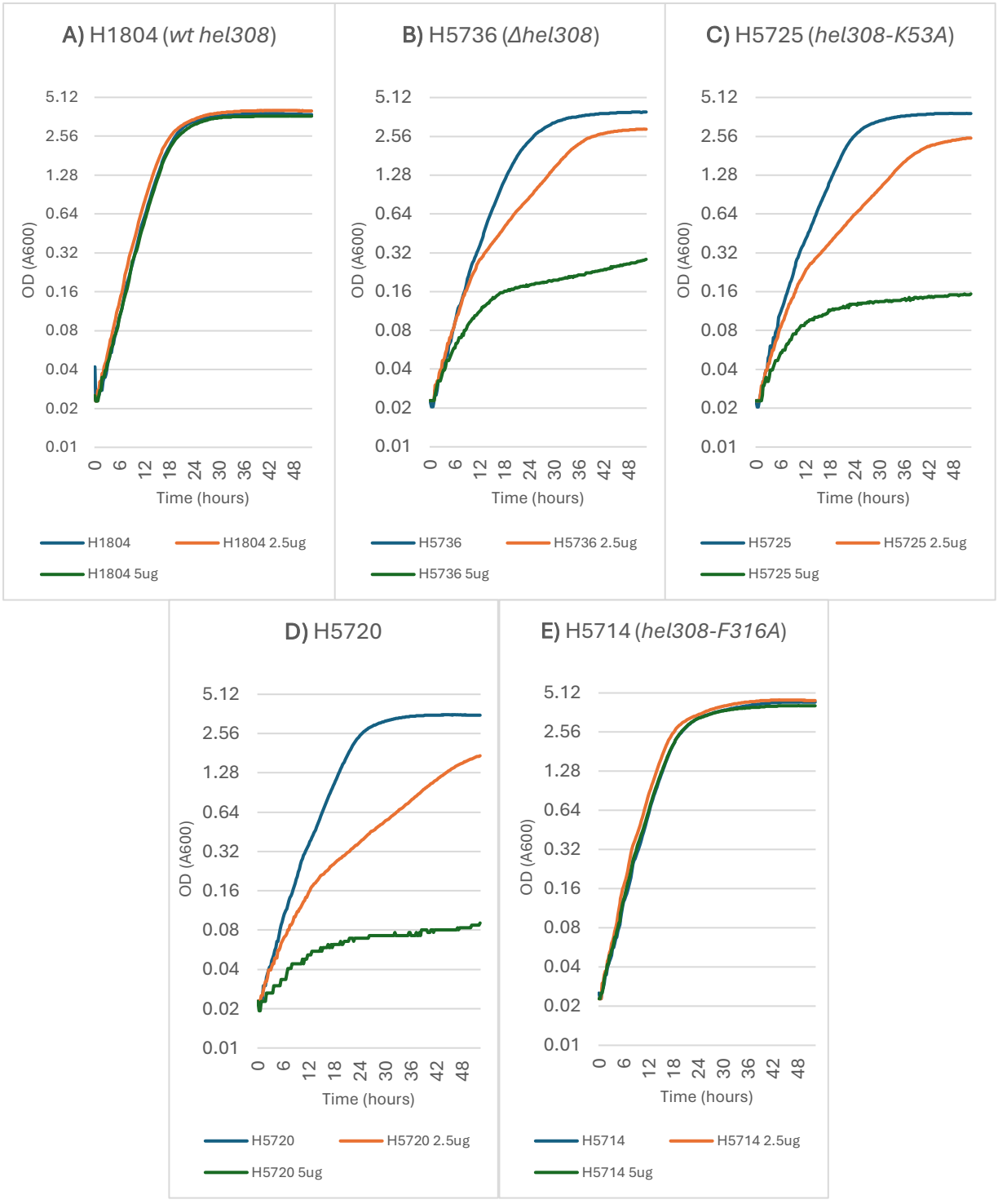


Figure 89. Effect of lower-dose aphidicolin on growth rate of originless strains, separated by strain.

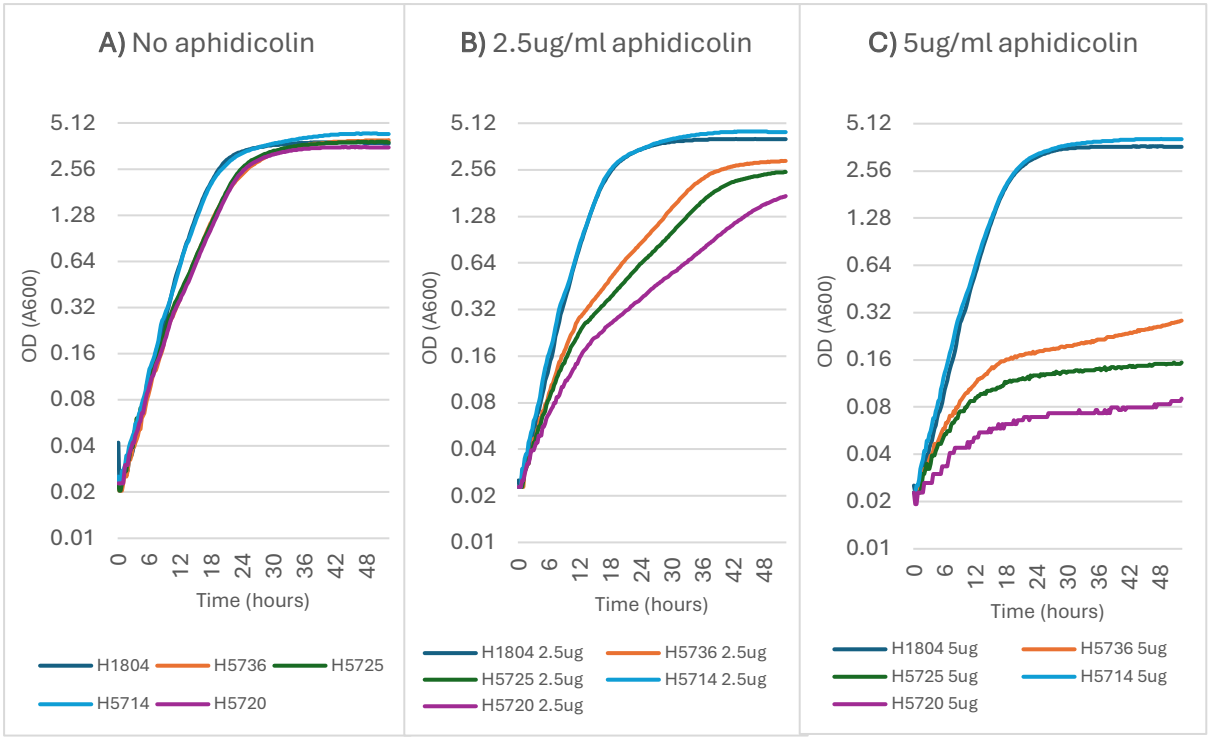


Figure 90. Effects of low-dose aphidicolin on growth rate of originless strains, separated by dose.

While it appears that H1804 and H5714 are barely affected by the lower doses of aphidicolin, the lower dose does allow some resolution of the responses of the *Δhel308*-like strains. In both the 2.5 and 5µg/ml conditions, H5736 (*Δhel308*) seems to have a small growth advantage over H5725 (*hel308-K53A*), which has a small growth advantage over H5714 (*hel308-D145N*). This seems to suggest that the presence of ATPase-null Hel308 is weakly detrimental to the cell; perhaps binding its substrates and blocking alternative mechanisms to resolve recombination intermediates and stalled forks. However, this effect may not be significant, as these data are produced by only three biological replicates for each condition.

As seen in chapter 4, there seems to be a weak biphasic growth pattern present in some of the strains, most evident in Figure 90B. The *Δhel308*-like strains seem to show a fast exponential phase, followed by a slower exponential phase, followed by a gradual transition into stationary phase. Had this not previously been observed in a previous chapter, in the absence of aphidicolin, it would be tempting to explain this as a delayed-onset or cumulative effect of aphidicolin treatment, whereby the chemical takes time to penetrate the cell and exert its effects. It is possible that this is evidence of some kind of quorum-sensing mechanism to co-ordinate growth rate within the population. Quorum sensing is not well-studied in archaea, but there is some evidence for the use of N-acyl homoserine lactones (AHLs) as a quorum sensing molecule across a range of halophilic archaea (Charlesworth et al., 2020).

It is also possible that this secondary, slower phase of growth could be brought about by the exhaustion of stores, or accumulation of aphidicolin effects within the cell. If the aphidicolin is interfering with genome replication, its effect may not be seen until a few generations have been produced; being partially masked by the higher ploidy during exponential phase growth. The cells may therefore be able to maintain the normal rate of division until the reduction in ploidy becomes a limiting factor. This could potentially be explored through the use of flow cytometry to quantify the DNA content (and thus inferred ploidy) of the cells at set time points following addition of aphidicolin. Unfortunately, this experiment could not be performed due to time constraints.

The more intriguing effect observed in these data is that, for the wild-type and F316A point mutant, 2.5µg/ml aphidicolin appears to confer a small benefit, resulting in the strains entering stationary phase at a higher OD. This is less obvious in the graphs with logarithmic Y axes, but was noticeable prior to this transformation. The non-logarithmic data for these two strains are shown in Figure 91, below.

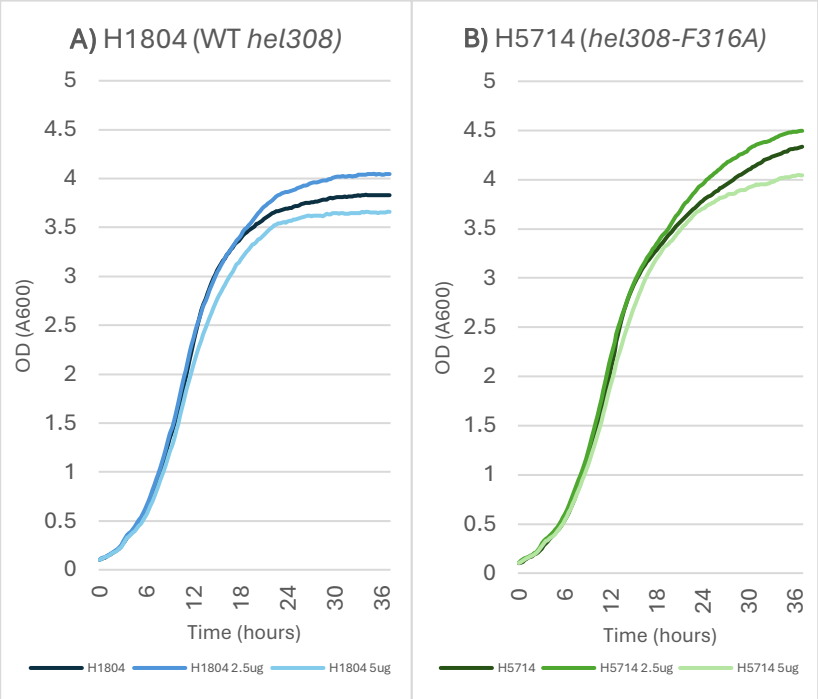


Figure 91. Growth curves for H1804 and H5714 under low doses of aphidicolin.

While the ODs did not increase by much in the 2.5µg/ml aphidicolin treatment, the fact that the effect was independently observed in both strains does support the legitimacy of the effect. It is not immediately clear why a small amount of PolB inhibition would be beneficial to the cells, especially as high concentrations of aphidicolin (12.5µg/ml) are still severely inhibitive of growth (Figure 88C). It could be that a low level of PolB inhibition reduces competition at replication forks, or liberates replisome components from failed PolB-initiated replication events for alternative replication pathways. It is also theoretically possible that higher doses of aphidicolin could be causing off-target effects within the cell.

6.4 Aphidicolin response in strains with wild-type origins

In order to better contextualise the above results from origin-less strains, the experiment was repeated in strains with wild-type origins. However, the three point mutants used above were not all available in matching strain backgrounds, and further genome editing could not be carried out due to time constraints. The strains used are therefore the following:

Table 6-2 Strains used in this section

Strain	Genotype	Comments
H164	*ΔpyrE2* *ΔtrpA* *bgaHa-Bb* *leuB-Ag1*	Produced by Thorsten Allers. Strain produced for recombination assays. Leucine auxotroph.
H4361	*ΔpyrE2* *ΔtrpA*	Produced by Beccy Lever. Descended from H164.

	bgaHa-Bb *leuB-Ag1* *Δhel308*	Leucine auxotroph.
H5721	*ΔpyrE2* *ΔtrpA* *hel308-K53A*	Produced for this study.
H2400	*ΔpyrE2* *ΔtrpA* *bgaHa-Bb* *leuB-Ag1* *hel308-D145N*	Produced by Rebecca Gamble-Milner. Descended from H164. Leucine auxotroph.
H2397	*ΔpyrE2* *ΔtrpA* *bgaHa-Bb* *leuB-Ag1* *hel308-F316A*	Produced by Rebecca Gamble-Milner. Descended from H164. Leucine auxotroph.

H5721 had been produced in a previous chapter to examine the effects of chromosomal replacement of *hel308* with *hel308-K53A*. It therefore does not possess exactly the same background as the other strains, which were produced to examine the effects of *hel308* variants in recombination assays. This does introduce a potential confounding factor into the results; namely, that the background of H5721 does not perfectly match that of the other strains used. It is possible that the presence of an intact *leuB* allele could confer a small growth advantage to this strain. However, the assay was conducted in YPC, which contains sufficient leucine for *ΔleuB* strains, so it is hoped that any advantage gained by H5721 would not be significant.

6.4.1 Response to aphidicolin

Responses to aphidicolin of the strains with wild-type origins are shown in Figures 92 and 93, below, separated first by strain and then by aphidicolin concentration. Three biological replicates were used per strain, per condition.

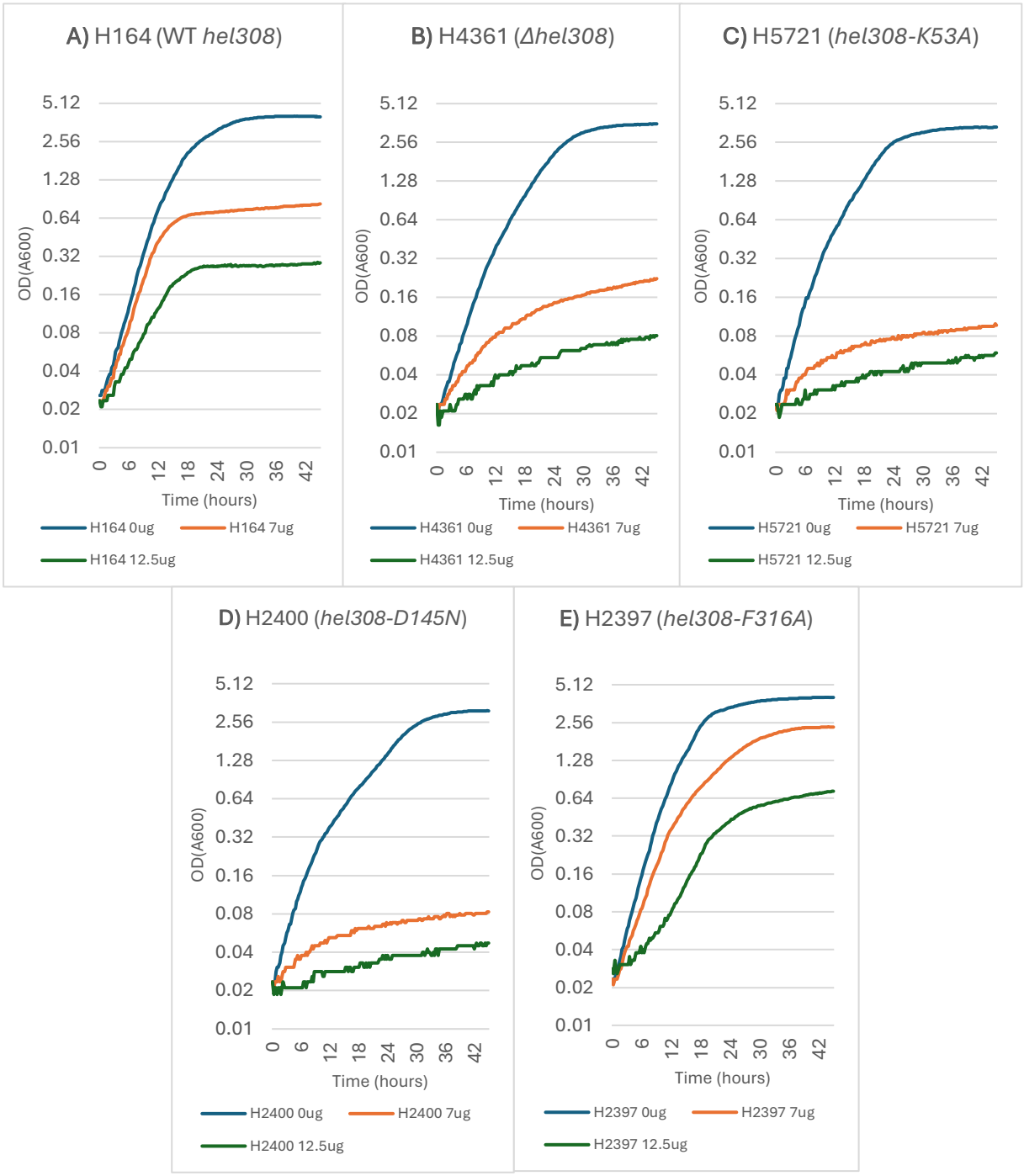


Figure 92. Effect of aphidicolin on strains with wild-type origins, separated by strain.

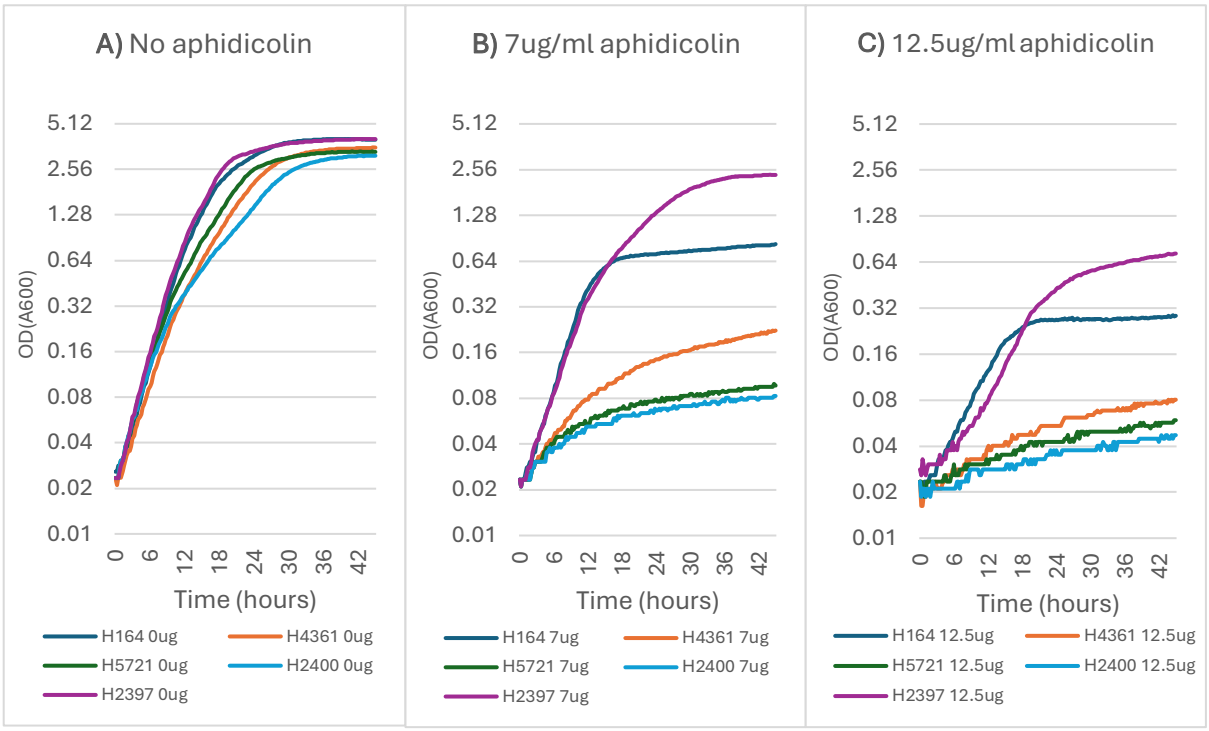


Figure 93. Effects of aphidicolin on strains with wild-type origins, separated by aphidicolin concentration.

As can be seen in Figure 92, the same two phenotypes are in evidence; a *hel308* wild-type-like, and a *Δhel308*-like. The ATPase-null point mutants behave similarly to the *Δhel308* strain, and may even show a slightly more severe defect under aphidicolin challenge, as was observed in the originless strains. Figure 93A also seems to indicate a greater difference in growth rate between the strains, suggesting that the absence of ATPase-active Hel308 is more detrimental in strains with intact origins than it is in the originless strains. The impact of the hyper-helicase point mutant F316A also seems to have greater effect in these strains, with H2397 entering stationary phase at a much higher OD than H164.

As was established by Smith (2021), the originless strains exhibit much lower aphidicolin sensitivity than the strains with intact origins. Under challenge with 7µg/ml aphidicolin, H1804 showed a generation time of around 2.25h, compared to 5.25h for H164. However, caution should be exercised when attempting to directly compare data from the two experiments, as the strains involved do possess other differences in their background, in addition to the deletion of origins from H1804.

6.4.2 Response to low-dose aphidicolin

To ensure comparability between the wild-type origin and origin-deleted strains, the effects of lower-dose aphidicolin (2.5 and 5µg/ml) were also investigated. The results are shown in Figures 94 and 95, below, separated first by strain and then by aphidicolin concentration. Three biological replicates were used per strain per condition.

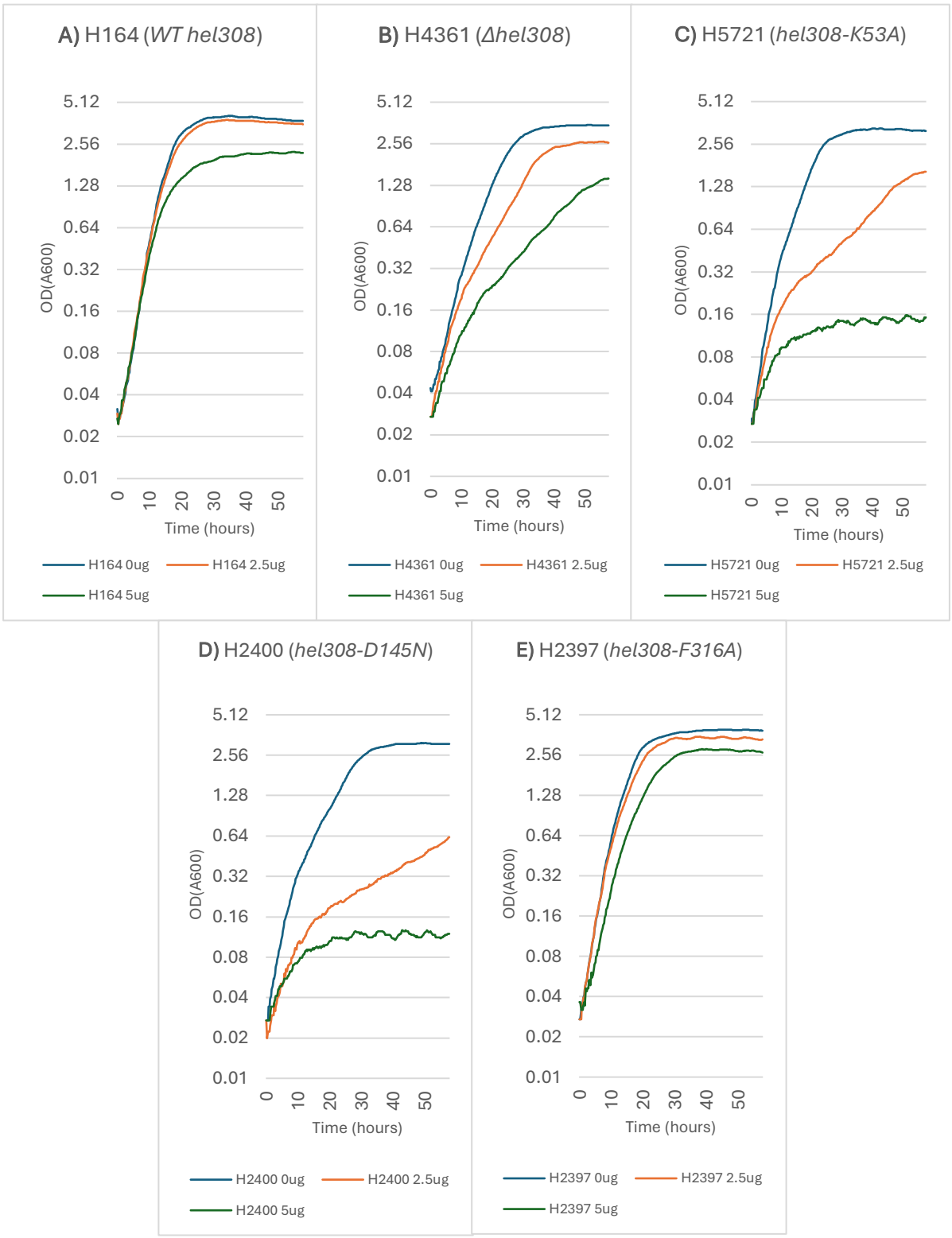


Figure 94. Effects of lower-dose aphidicolin on strains with wild-type origins, separated by strain.

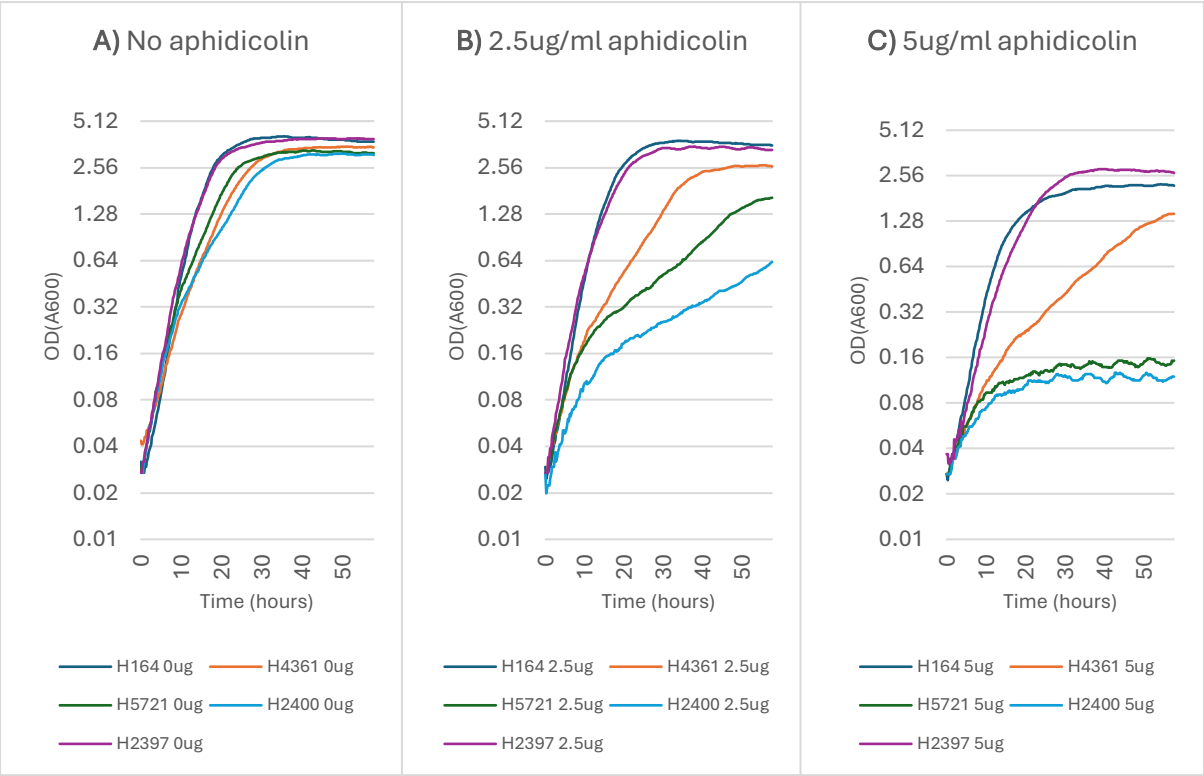


Figure 95. Effects of lower-dose aphidicolin on strains with wild-type origins, separated by aphidicolin concentration.

These effects appear broadly similar to the results from the originless strains – with the exception that there is seemingly no dose of aphidicolin which causes an increase in growth rate in strains with wild-type origins. The pattern of aphidicolin response is also broadly the same between *hel308* alleles in the two backgrounds – the *hel308-F316A* point mutant tolerates aphidicolin better than the wild type. The order of decreasing tolerance from *Δhel308* to *hel308-K53A* to *hel308-D145N* is also maintained as a pattern, offering some validation for this pattern in the above results. It is noticeable that, in Figure 94C, the two ATPase-null point mutants (H5721 and H2400) fare distinctly worse than the *Δhel308* strain. This could suggest that the helicase-inactive Hel308 is in some way obstructive within the cell. As previously mentioned, it may bind to its ssDNA targets, but then be unable to translocate, causing an obstruction to alternative pathways to resolve recombination intermediates and stalled forks.

These results also seem “noisier” than the previous assays – in some cases, a distinct sinusoidal effect can be seen, although at higher ODs this is somewhat mitigated by the logarithmic Y axis. As this effect has not been observed in any of the other assays with these strains, this may be attributed to an error in the machine. It seems to affect the numerical values of the measurements in a manner proportional to the OD – dense cultures seem to oscillate more than the cultures with severe growth defects. It is unknown what caused this error, but it has not been observed before or since.

6.5 Microdose aphidicolin effects

To explore the increase in growth rate observed in H1804 and H5714 in the presence of 2.5µg/ml aphidicolin, a range of low aphidicolin concentrations were investigated. The strains used in this investigation were H1804 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*) and H53 (*ΔpyrE2, ΔtrpA*) as a control. It was hoped that this could elucidate whether there was a low concentration of aphidicolin which would be beneficial to the growth rate of strains with wild-type origins.

6.5.1 Response to micro-dose aphidicolin

As it was already known that 5µg/ml caused a small growth defect in H1804 (see Figure 91A), a range of values between 0 and 5µg/ml were selected for this investigation. The results are shown below, separated by strain for clarity. Five biological replicates were used for each strain and dose combination. The results can be seen in Figure 96, below. Non-logarithmic Y axes have been used to make the small differences in maximum OD clearer.

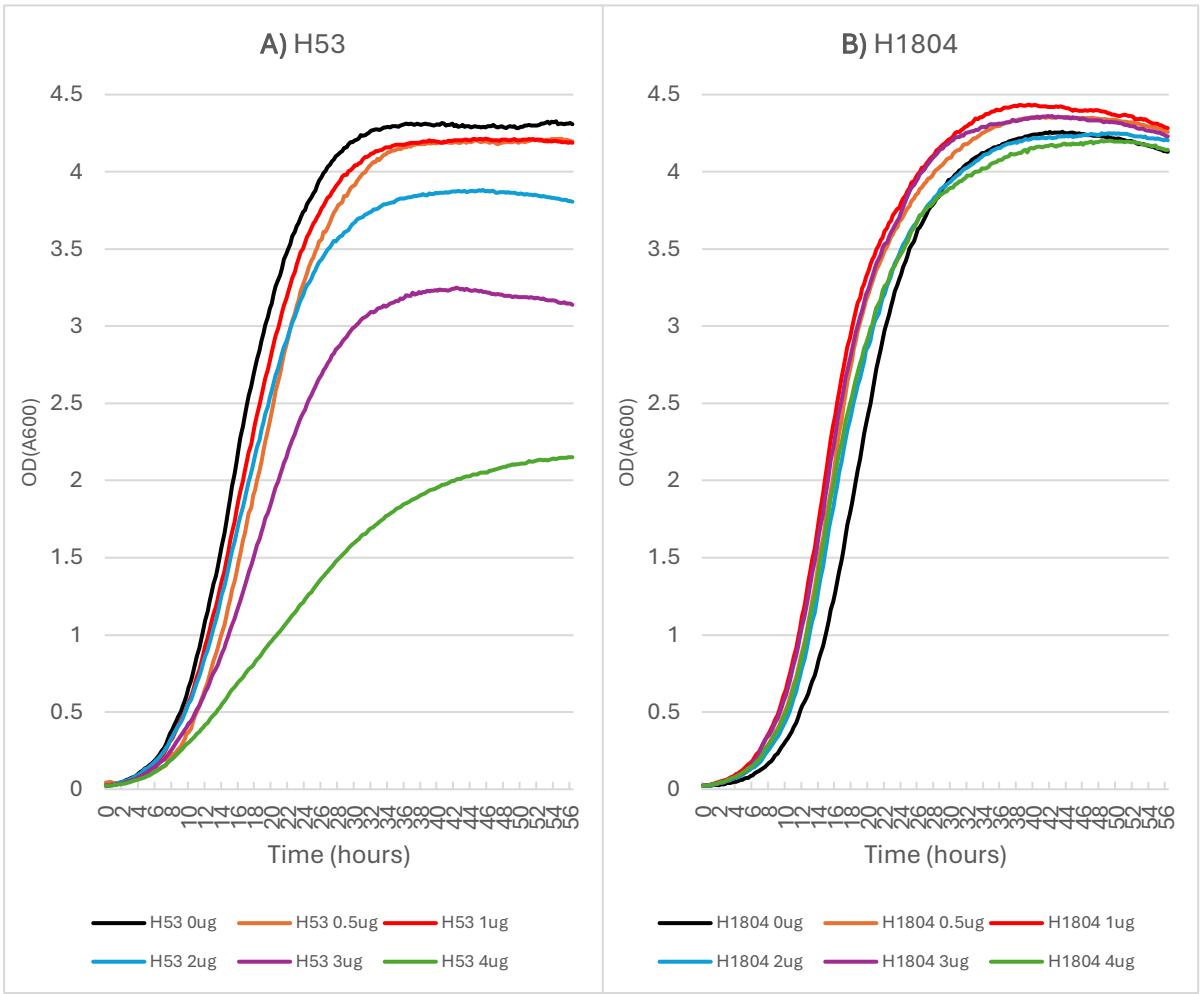


Figure 96. Effects of extremely low-dose aphidicolin on strains with intact origins (A) and chromosomal origins deleted (B).

It can be clearly seen that H53 demonstrates a dose-dependent inhibition of growth in the presence of aphidicolin. Even the lowest dose (0.5 µg/ml) causes a small but noticeable decrease in growth rate. By contrast, several of the treatments of H1804 can be seen exceeding the OD of their no-aphidicolin counterpart (shown in black). However, it is unusual that, in the H1804 data, the 3µg/ml growth condition exceeded the OD of the 2µg/ml. It seems unlikely that 1µg/ml and 3µg/ml should

both be beneficial, while 2µg/ml is less so, so it is tempting to attribute this to a pipetting error or insufficient mixing when adding the aphidicolin to the YPC broth. It is however apparent that no dosage of aphidicolin has been shown to increase the growth rate in strains with intact origins, while a range of low-aphidicolin treatments have been shown to be beneficial in strains lacking origins.

6.6 Discussion of responses to aphidicolin

When considering the interaction between origins, Hel308 and aphidicolin, it is difficult to provide an explanation which addresses all of the effects observed. However, a theory is presented below.

The most obvious explanation for the difference in aphidicolin susceptibility between originless and wild-type cells is that cells with wild-type origins are more reliant on PolB for genome replication, as evidenced by their increased susceptibility to aphidicolin. This is in line with studies in *T. kodakarensis*, which have shown that PolB is inessential in this organism, which exclusively performs origin-independent replication (Gehring et al., 2017). The *T. kodakarensis* data imply that PolB is the dominant DNA polymerase used for origin-dependent replication, while PolD is the main DNA polymerase used for recombination-dependent replication.

However, originless *H. volcanii* cells do not possess any additional factors or proteins that wild-type cells lack. At first glance, it appears that any cellular activity undertaken by originless cells should be performed, with equal ability, by the wild-type cells. Should replication via PolB be blocked, why can they not simply revert to the recombination-dependent replication method which appears to be undertaken in originless cells? Similarly, why would Hel308, a helicase which reduces the rate of recombination (Lever, 2020), be important to the survival of recombination-dependent originless cells when challenged with aphidicolin? It could be that the action of Hel308 is somehow integral to the replication strategy employed by cells deprived of active PolB. However, I believe that the small growth increase observed in H1804 and H5714 is a crucial piece of information in explaining the relationship between these factors.

Aphidicolin inhibits PolB by entering its active site during DNA polymerase activity (Baranovskiy et al., 2014). While this interaction is transient and does not cause any chemical changes to either component, this is sufficient to stall the replication fork, posing a barrier to further replication, as well as other DNA metabolism and maintenance activities (Vesela et al., 2017). As a further consequence, studies in eukaryotic systems have shown that, following aphidicolin treatment, the replicative helicase can uncouple from the DNA polymerase (Pacek et al., 2006). This uncoupled helicase (as part of the CMG complex) rapidly progresses unwinding of the replication fork, leaving long stretches of vulnerable ssDNA (Byun et al., 2005). Aphidicolin treatment of cells undergoing DNA synthesis via PolB will therefore result in multiple stalled replication forks.

Archaeal Hel308 has been shown to target forked DNA structures *in vitro*, and may be more active at processing these structures than it is at migrating Holliday junctions (Guy & Bolt, 2005). Combined with its observed ability to displace proteins from DNA, it has been suggested that this protein plays an important role in repairing stalled replication forks (Richards et al., 2008). In both of these studies, a preference for unwinding of the lagging strand was observed.

Stalled forks can be repaired by recombination, in which Hel308 is thought to participate in resolution of Holliday junctions through migration of forks and an association with resolvase Hjc (Hong et al., 2012). The long ssDNA strands created by uncoupling of the replicative helicase from the DNA polymerase (Byun et al., 2005) could also increase frequency of DNA breaks, which could

become sites of break-induced repair. Low levels of aphidicolin-induced inhibition of PolB therefore result in an increase in initiation of recombination events to repair the stalled forks. As origin-less strains are reliant upon recombination for replication of their genomes (Hawkins et al., 2013), this small increase in recombination events may be beneficial to the cell by increasing opportunities for (or efficiency of) genome replication. However, at higher levels of aphidicolin, the burden of removing so many stalled forks may outweigh the benefit of the increased recombination frequency.

This would explain the behaviour of the originless strains in the presence of aphidicolin, wherein small doses of aphidicolin are beneficial, but greater doses result in reduced growth rate. In strains lacking functional Hel308, stalled forks may be repaired through another, less efficient pathway, or perhaps are not converted to opportunistic replication events as effectively. This theory does of course presuppose that PolB is still active at a low level in originless cells – which is likely, given the essentiality of the *polB1* gene even in cells without chromosomal origins (Smith, 2021). In the strains used above, the origins on the minichromosomes are still present, and may be replicated by PolB.

It could be that, if there is greater PolB activity in strains where origins are present, there is more potential for damage and stress to be caused where these origin-initiated replication forks are stalled. During exponential phase, ploidy in *H. volcanii* is around 30 (Zerulla et al., 2014). To support a doubling time of 2.5h (which is approximately that of wild-type *H. volcanii*), all 30 copies must be replicated at least once, on average. Supposing that all six origins (four on the chromosome and one on each minichromosome) fire on average once per cycle ((Hawkins et al., 2013) showed that some origins are more active than this, being represented at around 2.2 times more than the least represented regions – but an average of once per origin per cell cycle could be reasonable); each origin activates bidirectionally, and so recruits four DNA polymerases (two per replication fork and two forks per origin) – presumably PolBs, as this is thought to be the dominant replicative polymerase. This results in recruitment of 30x6x4 = 720 PolBs recruited in each 2.5h window. Only one molecule of aphidicolin is needed to stall a single PolB enzyme (Baranovskiy et al., 2014). In the presence of even low concentrations of aphidicolin, strains with active origins could easily generate a huge number of stalled replication forks, causing replicative stress. As neither aphidicolin nor PolB are chemically altered or broken down as a result of their interaction (Baranovskiy et al., 2014), it would be possible for the same two molecules to produce multiple stalled fork events, should the stalled fork be reset. Even if PolB's role is limited to either the leading or lagging strand in origin-initiated replication forks, this is still potentially a huge number of stalled fork events for the wild-type cells to repair.

While much of this is speculation, some limited support may be found in human cellular studies – Chaudhury et al. (2014) used aphidicolin to induce stalled forks, and observed that FANCD2 and BLM helicase (a superfamily 2 helicase) were important factors involved in their repair. In yeast, Pif1 helicase plays an important role in initiation of break-induced replication (Li et al., 2021). It is possible that Hel308 is playing an equivalent role to these helicases in *H. volcanii*. However, as recombination and break-induced replication are effectively primed by broken DNA strands rather than RNA primers, activity of PolD at these sites potentially contradicts reports from *T. kodakarensis* and *P. abyssi* that PolD is primarily an extender of RNA-primed strands rather than DNA-primed strands (Greenough et al., 2015; Henneke et al., 2005). Studies in *P. abyssi*, however, have shown that both B and D polymerases are able to perform recombination-primed DNA replication (Hogrel et al., 2020). The preferences of substrate of the DNA polymerases in *H. volcanii* are as yet uncharacterised.

6.7 Aphidicolin response in *H. mediterranei*

The mechanism or factors that render *H. mediterranei* incapable of origin-independent replication have not yet been identified. As it seems the case that PolD is the dominant DNA polymerase in the absence of origins, it is possible that there is some biochemical difference between the PolD and PolB homologues in the two species. In order to explore this, two *H. mediterranei* strains were selected and subjected to the same treatments described above. As origins are essential in this species, it was not possible to test the activity of an origin-less strain; the strains chosen were instead H828 (*ΔpyrE2, ΔtrpA*) and H4676 (*ΔpyrE2, ΔtrpA, ΔoriC1, ΔoriC2, ΔoriC3*). In H4676, the three canonical chromosomal origins have been deleted, leaving only the usually-dormant origin, OriC4, to replicate the chromosome (Yang et al., 2015). As the deletion of all chromosomal origins is not possible in this species, it was thought that this strain was as close to H1804 as could be easily replicated in this species. As with H1804, the minichromosomes and their origins are still present.

The two *H. mediterranei* strains were trialled against low concentrations of aphidicolin, so that their behaviour could be compared to the *H. volcanii* strains in this regard. It was expected that these would cause growth defects in both *H. mediterranei* strains, given their inability to engage in origin-independent replication. The data are shown in Figure 97. Data shown are the result of five biological replicates per strain per condition.

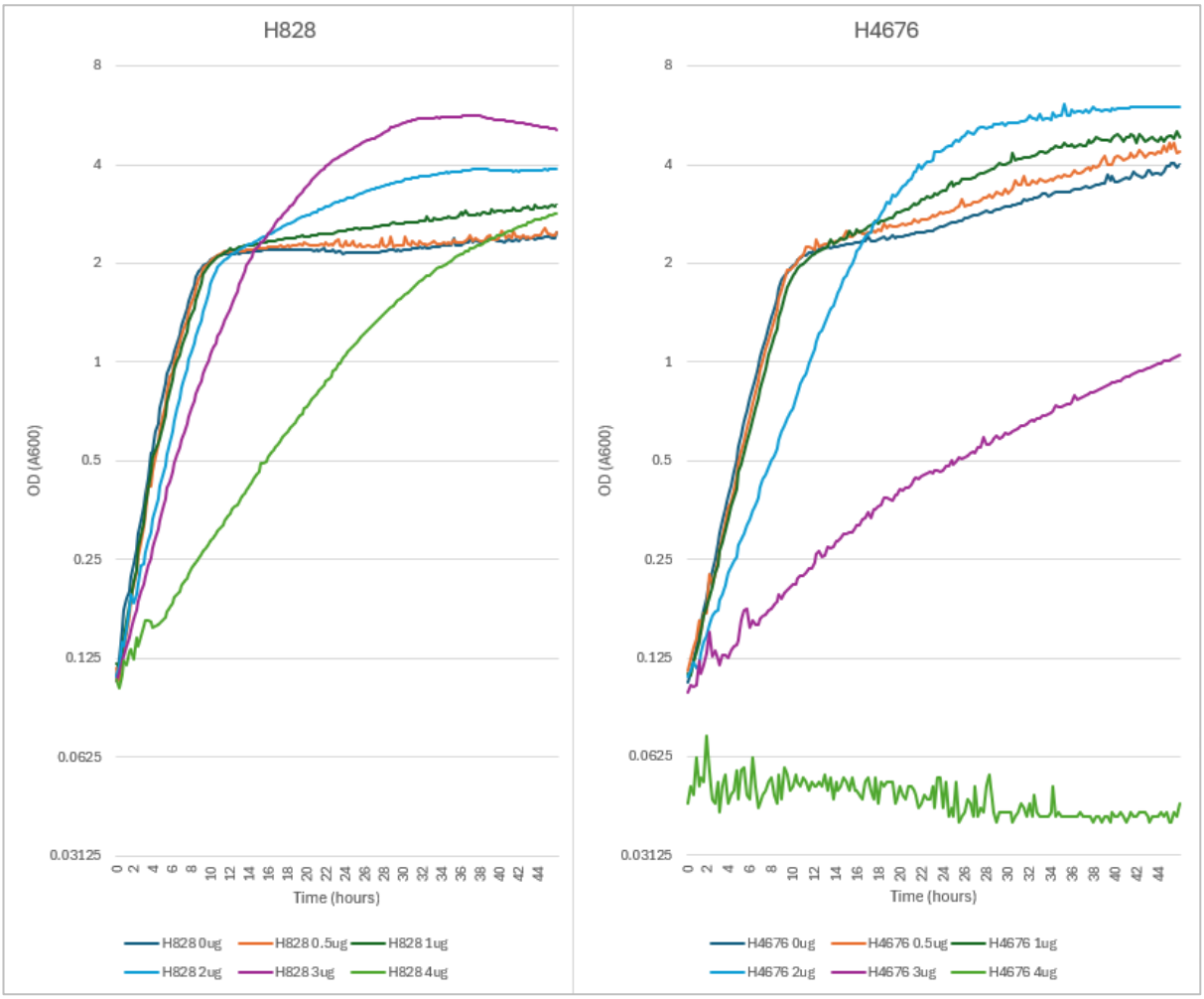


Figure 97. Growth of H. mediterranei *strains H828 and H4676 challenged with microdose aphidicolin.*

To our knowledge, this assay has not been attempted in this species before; consequently it is difficult to interpret the data, as it is hard to know what to expect. A search of the literature identified no published studies showing a "normal" growth curve for this species. The data is also shown in Figure 98, below, with a linear y-axis, to allow comparison with the data for *H. volcanii* strains under the same treatment.

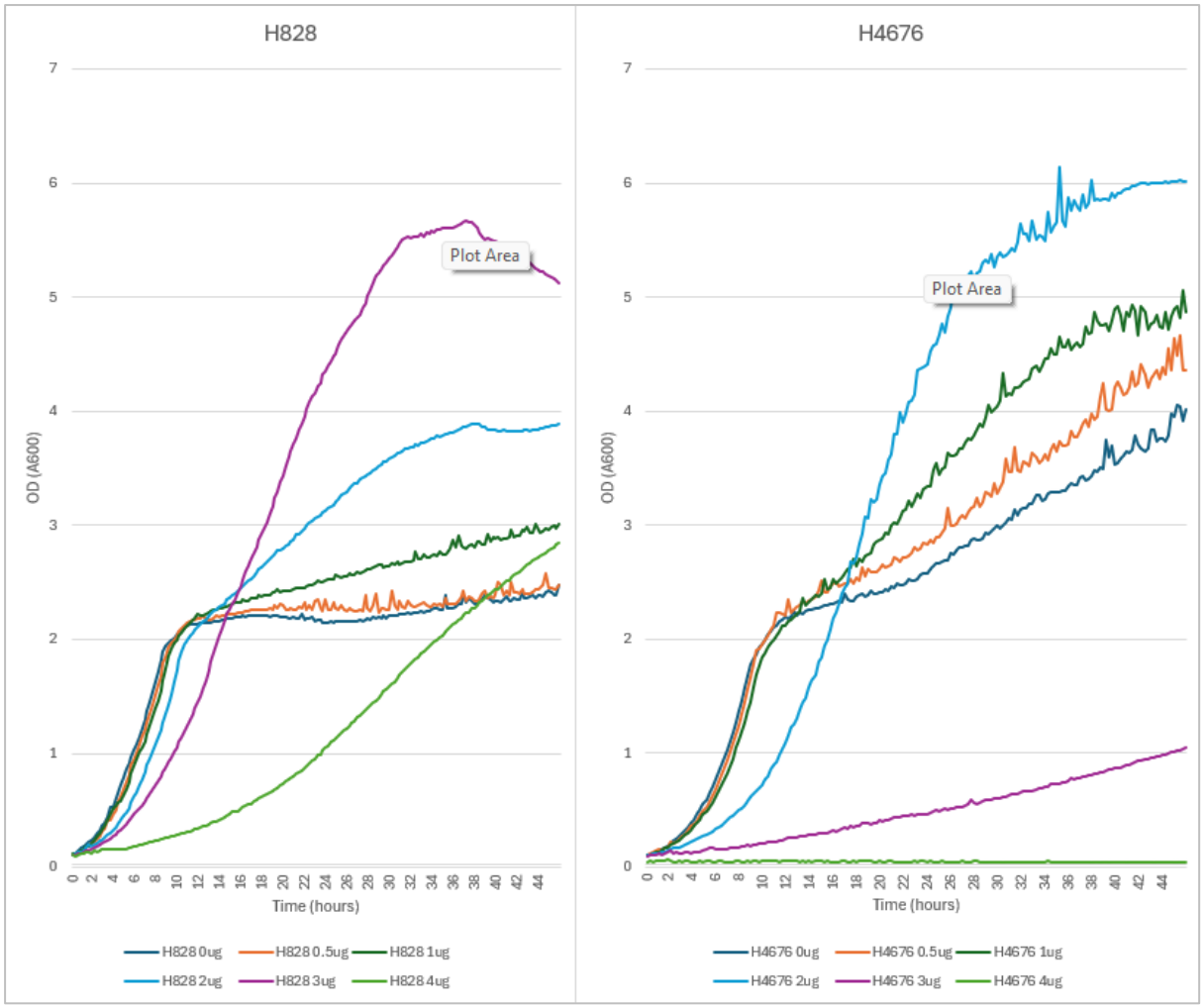


Figure 98. Growth of H. mediterranei *strains H828 and H4676 challenged with microdose aphidicolin.*

Under different treatments, the two strains seem to undergo very different growth patterns. 0.5 and 1µg/ml of aphidicolin seem to induce very little change in the behaviour of either species; however, 2-3µg/ml seems to inhibit growth rate, but also to increase the maximum OD reached considerably.

It is possible that the growth rate shown in the no-aphidicolin treatment shows biphasic growth, or some kind of quorum sensing mechanism. It is noticeable that, although the cells used for each strain were part of the same culture before the start of the assay, they begin behaving very differently in response to only small changes in aphidicolin dose. Under no-aphidicolin conditions, the two strains seem to reach stationary phase at around the same OD, although H4676, having a longer doubling time, as reported by Yang et al. (2015), takes longer to reach this point. H4676 however continues slowly growing after this point is reached, while H828 remains mostly stable, growing much more slowly. The sharpness of the transition between exponential and stationary phase is also much greater than in *H. volcanii*. However, it seems that at certain doses of aphidicolin, this sharp plateau may be avoided.

The inhibition seen at 4µg/ml aphidicolin is much higher in this species than it is in *H. volcanii* strain H53, above. It could be that this indicates a much higher reliance on PolB in *H. mediterranei*, or it could be that mechanisms of stalled fork repair are much less effective in this species. Unlike *H. volcanii*, in *H. mediterranei* the strain with fewer origins seems to exhibit much greater inhibition in response to aphidicolin treatment. It could be that, given the inability of this organism to undergo

origin-independent replication, the sole origin present in H4676 represents a significant point of failure. H828, possessing all three wild-type chromosomal origins (plus one dormant origin) has some redundancy, whereby failure of one fork initiated by the origins can be compensated by forks progressing from the other origins. In H4676 however, the single origin is responsible for replication of the entire genome; disruption of replication may therefore confer a greater growth defect due to a lack of redundancy.

As this method of growth assay has not been validated for use in this organism, it is possible that the strange growth curves seen in Figures 97 and 98 are not accurate reflections of the density of the cultures. *H. mediterranei* is known to produce extracellular polysaccharides (as described in Antón et al. (1988)), which are the cause of the "slime" observed when this species is grown on solid media. It is possible that these polysaccharides could result in an increased viscosity of the liquid in the wells, leading to uneven density of cells, or otherwise interfering with penetration of light through the culture.

While it would be interesting to observe the response of *H. mediterranei* to the other doses of aphidicolin trialled above, this experiment could not be completed within the time frame, being stymied by machine error on three separate occasions. While the results gathered in this subsection are certainly intriguing, they are difficult to comment on given the absence of other data or literature sources to refer to.

6.8 Future work

6.8.1 Quantification of growth promotion/inhibition by competition assay

For some of the aphidicolin assays shown above, it is clear to see that growth rate is altered, as indicated by a difference in gradient during the log phase of growth. However, in the instances where growth rate is similar between different strains, it is difficult to say with certainty which strain has a growth advantage, and by how much. More quantifiable data could be generated through competition assays, as these make differences in growth rate much more obvious by allowing direct comparison between strains. In many cases, competition assays can show comparative differences in growth rate that would be extremely difficult to observe or measure by growth assay.

This approach would require the addition of a marker gene to allow identification of the two different strains on solid media. Markers that are not expected to confer a selective advantage are often used for this purpose, such as mutations to pigment genes, or inclusion the *bgaHa* gene for blue-white screening. Growing different pairs of strains in the presence and absence of aphidicolin, and plating samples from the culture at intervals to visualise the relative populations would allow measurement of the growth rates of the strains in direct comparison with each other.

This series of experiments could not be attempted during this project due to time constraints.

6.8.2 Response to aphidicolin in different strain backgrounds

While it is assumed that most or all of the effects of aphidicolin inhibition in *H. volcanii* is due to inhibition of PolB1, this has not yet been conclusively proven. *H. volcanii* contains two PolB genes, and while PolB2 is not thought to play an integral role in genome replication, it may be playing some role in the cell, potentially related to DNA repair or translesion synthesis, as seen in *S. islandicus* (Feng et al., 2022). As this gene is inessential in *H. volcanii* (T. Allers, unpublished research), the effects of its deletion could be explored, and the response of strains lacking this protein compared to those with both paralogues.

The presence of the two minichromosomes, particularly pHV3, in the "originless" strains used to test aphidicolin response, is potentially a confounding factor, as these two minichromosomes still possess origins of replication, and presumably replicate using the same mechanisms as the wild-type chromosomes. In order to explore the influence of these origins, they could both be integrated onto the main chromosome and then have their origins deleted, to produce a truly originless strain. The response of this strain to aphidicolin could then be characterised.

6.8.3 Examination of replication profiles in aphidicolin-treated cells

Replication profiles are a powerful tool for examining where the common sites of replication are originating. Hawkins et al. (2013) previously used this data to determine the capability of *H. volcanii* to undergo origin-independent replication, and Gehring et al. (2017) did likewise to prove that *Thermococcus kodakarensis* does not use its origins even when they are intact. It is possible that this technique could also be used to elucidate the roles of the DNA polymerases in *Haloferax.*

Comparison of replication profiles from strains with origins before and after aphidicolin treatment could provide data on DNA polymerase usage. In wild-type *H. volcanii,* replication profiles show clear peaks corresponding to the origins, as shown in Figure 99, below.

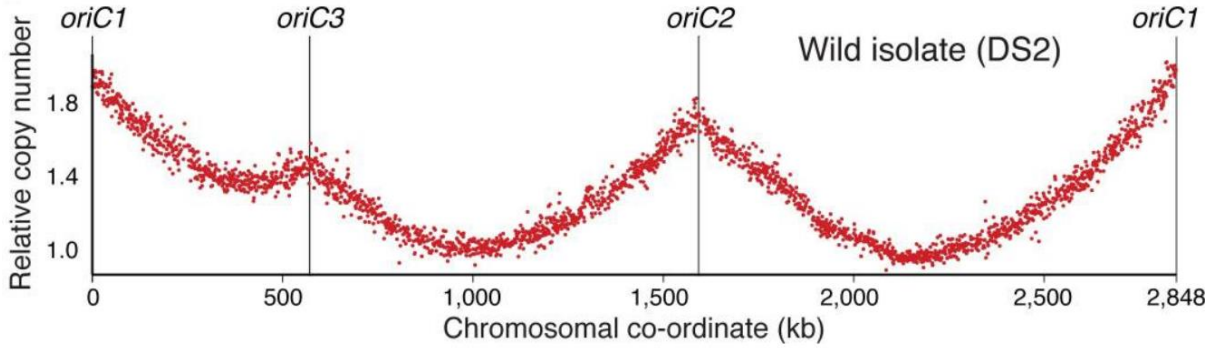


Figure 99. Replication profile of wild-type H. volcanii. *Data from Hawkins et al. (2013).*

If PolB is the main replicative DNA polymerase in *H. volcanii*, treatment with aphidicolin should result in a flattening of the profile, as inhibition of PolB results in alternative methods of genome replication. Unfortunately, this technique cannot shed light on PolB usage in originless strains, as the replication profile of these is already mostly flat, as seen in Figure 100, below.

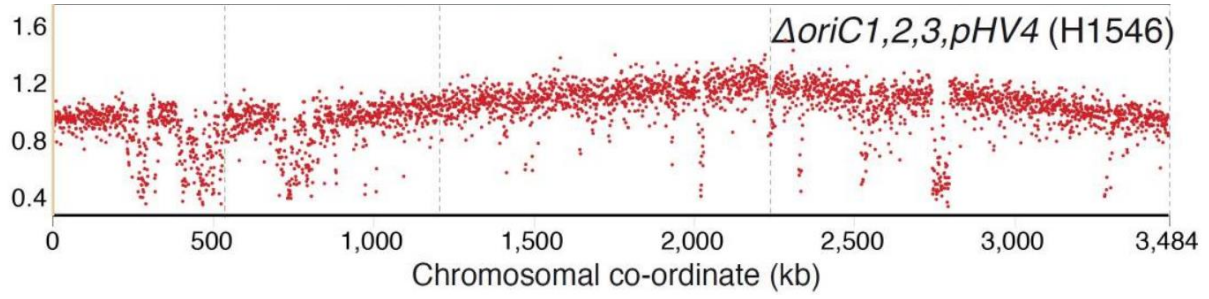


Figure 100. Replication profile of an H. volcanii *strain in which all chromosomal origins have been deleted. Data from Hawkins et al. (2013).*

This experiment in particular could add a lot of evidence as to whether PolB is the canonical DNA polymerase during origin-initiated genome replication. However, it is possible that, if stalled forks are important sites of *ad-hoc* replication under aphidicolin treatment, significant peaks in the replication profile may still be visible near the origins, as a result of attempted origin-initiated DNA replication, followed by fork stalling and opportunistic recombination-based replication.

6.8.4 Comparison of aphidicolin-treated cells with reduced-PolB cells

As discussed above, it is possible that some of the effects observed following aphidicolin treatment are not solely due to reduced activity of PolB, but by it either stalling or dissociating during DNA replication. Perhaps the two effects can be examined separately by controlling the levels of this protein within the cell.

H. volcanii possesses two *polB* genes; *polB1* and *polB2*. Of these, *polB1* is essential, while *polB2* is inessential. Given its inessentiality, it is considered unlikely for *polB2* to be playing an integral role in replication. The PolB2 family is usually inactive as a DNA polymerase in the Archaea, as a result of mutations disrupting essential catalytic residues and abolishing exonuclease activity (Rogozin et al., 2008). This gene inessential in *H. volcanii*, and has been successfully deleted (T Allers, unpublished data).

Introduction of an inducible promoter - either *p.tnaA* or *p.tna*M3, as appropriate to wild-type expression levels - upstream of the *polB1* gene could allow some control over the expression levels of this protein. Expression would have to be controlled carefully, as this gene is essential. This would allow examination of any differences between PolB inhibition by aphidicolin and deprivation of PolB due to low expression. For example, if the increased growth rate observed in originless strains in response to low-dose aphidicolin is due to increased recombination events, a small decrease in PolB levels in this strain background should not induce any growth advantage. Likewise, the effects of *Δhel308* should be less severe in PolB deprived treatments versus PolB inhibited treatments, as there should be no increased burden of stalled forks to remedy.

While this is theoretically a sound experimental approach, previous attempts to place *polB1* under inducible control have not met with success (Smith, 2021). However, it is possible that complementation with inducible *polB1* from a plasmid could allow deletion of the chromosomal copy of the gene, leaving only the episomal, inducible copy in the cell.

Should introduction of inducible expression to the *polB1* gene not prove possible, overexpression of this protein may prove an alternative pathway to exploration of this phenomena. Addition of inducible *polB1* on an episome could allow determination of what proportion of the growth defect seen in *Δhel308* strains in the presence of aphidicolin is due to shortage of PolB, and how much is due to the burden of repairing stalled forks. If the majority of the growth defect is due to a shortage of PolB for replication activities, then the increased expression of this protein should confer some resistance to aphidicolin treatment. However, if the growth defect is due to difficulty in repairing stalled forks, the sensitivity to aphidicolin should remain unchanged. As aphidicolin inhibits PolB during the process of DNA polymerisation, the number of stalled forks produced would be proportional to the number of PolB enzymes engaged in DNA polymerisation activity, not the total number of this protein present in the cell.

Alternatively, knockdown of the *polB* gene could be attempted through CRISPR interference. CRISPR has been shown to be functional in *Haloferax* (Maier et al., 2019). It has also previously been shown to be capable of knocking down even essential genes to as low as 22% (Stachler & Marchfelder,

2016). Although previous attempts to silence *H. volcanii*'s *polB* using CRISPR interference have not met with success (Smith, 2021), there is still potential in this method to distinguish between aphidicolin-induced growth defects due to PolB deficiency, and aphidicolin-induced growth defects due to widespread replication fork stalling.

6.8.5 Independent control of origin activity and *hel308* expression

For a long time in the story of *H. volcanii* as a model organism, only one inducible promoter has been available; *p.tnaA* (Large et al., 2007). However, identification of a second inducible promoter (*p.xyl*, (Rados et al., 2023)) was published last year. For the first time, this could allow independent control of the expression of two separate genes in *H. volcanii*, by placing two different genes of interest under two different promoters.

The strain H5107 (described in the "screen verification" chapter) is a strain in which all origins were deleted, including those on the minichromosomes. One origin was then re-introduced, and placed under inducible control, allowing the strain to be manually switched between origin-dependent and origin-independent replication. *hel308* (or other genes of interest) could be placed downstream of the *p.xyl* promoter in this strain, allowing phenotyping of their behaviour in different combinations of conditions to be explored. This could allow direct comparison of the responses of strains expressing different levels and combinations of the relevant proteins to aphidicolin. It could also allow exploration of the effects of varying the levels of Hel308 within the cell, which was not explored during this study.

6.9 Conclusion

It was previously reported that deletion of origins causes reduced susceptibility to PolB inhibitor aphidicolin in *H. volcanii.* It was also reported that deletion of *hel308* causes susceptibility to aphidicolin in this species. The effects of both of these factors were explored through a series of aphidicolin assays, and it was concluded that a lack of active Hel308 produces a dominant aphidicolin-susceptibility phenotype.

While the exact mechanisms underlying this effect have not been conclusively proven, this provides some evidence for the role of PolB as the dominant DNA polymerase in origin-dependent replication, and a reduced activity of this enzyme in origin-independent replication. It is also suggested that Hel308 is an important factor in processing of stalled replication forks as a result of aphidicolin-based PolB inhibition.

Preliminary data on the response of *H. mediterranei* to aphidicolin treatment is intriguing, but difficult to interpret in the absence of other data and literature sources.

7. Conclusion

In the course of this work, several aspects of genome replication in the two *Haloferax* species have been explored, although few points have been conclusively confirmed. The findings are summarised briefly below.

Previous reports that the *H. mediterranei* Hel308 protein inhibited origin-independent replication in *H volcanii* could not be confirmed or replicated. Based on the experiments described in Chapter 3, it was determined that *H. mediterranei hel308* is sufficient to complement the growth defect caused by *hel308* deletion in *H. volcanii*, further corroborating the similarity of the two genes. However, it can be stated with confidence that the *H. mediterranei hel308* gene is insufficient for inhibition of OIR in *H. volcanii*. It should be noted that this does not preclude the possibility that this gene is an important factor in the inability of *H. mediterranei* to perform origin-independent replication. It is possible that *hel308* is only one of several necessary factors whose interactions inhibit this activity in *H. mediterranei*, or enable it in *H. volcanii*. Given the lack of conclusive data on this subject, several alternative approaches to elucidating this phenomenon were suggested in this chapter, which may provide an indication of further avenues to explore.

Previous reports of the lethality of the ATPase-null point mutant *hel308-K53A* could likewise not be corroborated. This lack of lethality allowed further exploration of this point mutant, and it was engineered onto the chromosome to allow some phenotyping assays, as described in Chapter 4. Strains bearing this point mutant demonstrated increased susceptibility to UV damage, similar to hel308-deleted strains and the alternative ATPase-null point mutant, D145N. Similarly, under normal growth conditions, both ATPase-null point mutants exhibit a growth defect similar to *Δhel308* strains. However, the three strains (*Δhel308, hel308-D145N* and *hel308-K53A*) are not perfectly equivalent; *Δhel308* growth slightly faster than the two point mutants, and *hel308-K53A* consistently grows slightly faster than *hel308-D145N.* This pattern also holds true when challenged with aphidicolin treatment, both in the originless strains, and those with WT origins (described in Chapter 6). This implies that ATPase-null *hel308* phenotype is not perfectly equivalent to the *hel308*-deletion phenotype; suggesting that there may be some activity or interaction that the enzyme is involved with that does not require ATP. Furthermore, the fact that Hel308-D145N (which is expected to be unable to bind ATP) and Hel308-K53A (which is expected to be able to bind, but not hydrolyse ATP) produce slightly different phenotypes could suggest that ATP binding alters conformation or activity of the enzyme. Future exploration could include exploring the effects of the *hel308-K53A* allele on recombination rate or response to further DNA damage agents such as MMC or MMS.

Reports of the essentiality of *hel308* in *H. mediterranei* were cast into doubt by the revelation that the deletion construct used in previous experiments inadvertently deleted the start codon of the *cgi* gene, which is known to be essential in many species. Attempts to delete *hel308* from *H. mediterranei* with an alternative deletion vector, as described in Chapter 5, did not bear fruit. Based on this, it seems likely that *hel308* is essential in *H. mediterranei*, or else its deletion confers a growth defect so severe as to make pop-outs very poorly represented in culture. Similarly, attempts to delete *mrr* from *H. mediterranei* did not result in the isolation of any homoploid deletion mutants. Further attempts to delete *hel308* under complementation from an episomal copy of the gene may clarify the essentiality of *hel308* in this species by limiting any growth defect in homoploid deletion mutants.

Chapter 6 covers a series of experiments exploring the interplay between the aphidicolin resistance incurred by chromosomal origin deletion, and aphidicolin susceptibility conferred by deletion of *hel308* in *H. volcanii*. *H. volcanii* contains both PolB and PolD family DNA polymerases, whose

preferred roles in DNA replication and repair have not been fully characterised in this species. Aphidicolin specifically inhibits the DNA polymerase PolB; use of this agent was intended to generate data on the comparative usage of different polymerase families in *H. volcanii*. It was found that deletion of *hel308* from strains lacking chromosomal origins caused increased susceptibility to aphidicolin. A similar phenotype was observed in strains where *hel308* had been replaced with ATPase-null *hel308* point mutants. In addition, the *hel308-F316A* point mutant, previously characterised as a hyper-helicase mutant, did not confer aphidicolin susceptibility. In fact, this point mutant seemed to confer a small increase in growth rate. In addition, during an experiment testing very low doses of aphidicolin in these strains, it was observed that extremely low doses of aphidicolin resulted in a small growth increase in the strains bearing *wt-hel308* and *hel308-F316A*.

Furthermore, equivalent experiments in strains possessing wild-type origins showed that deletion of *hel308* in this background also conferred increased aphidicolin susceptibility. Likewise, the same *hel308* point mutants showed similar phenotypes as in the originless background. However, there was no dosage of aphidicolin which improved the growth rate in strains with wild-type chromosomal origins. It should be noted that data from the originless and wild-type-origin strains are not directly comparable, owing to some minor differences in the genetic backgrounds used in strain construction.

Finally, aphidicolin response was trialled in *H. mediterranei* strains; one in which all origins are present, and one in which all but one had been deleted. Data from this experiment was hard to analyse, owing to a lack of previous literature to provide context.

Prior to this work, it had been assumed that the aphidicolin resistance seen in originless strains was due to a reduced reliance on PolB in this background; suggesting that perhaps PolB was used in origin-initiated replication, while PolD was the dominant DNA polymerase used in recombination-dependent replication. However, this model by itself does not explain why a small dose of aphidicolin would be beneficial in cells lacking origins.

The data generated in this study seems to suggest that *hel308* is an important part of aphidicolin response in both backgrounds, and that the role played by this enzyme specifically requires its ATP-dependent helicase activity. While the aphidicolin experiments were intended to explore the effects of PolB deprivation in the cells, it is suggested here that some of the effects observed may be due not to a lack of PolB, but due to a high prevalence of stalled replication forks where PolB is inactivated by aphidicolin. It is suggested that Hel308 is important in repair or recovery of these stalled forks, as has been previously reported for this enzyme. This explains why *Δhel308* strains are more severely affected by aphidicolin in both wild-type and originless backgrounds; an important pathway for stalled fork repair is damaged in these strains. The presence of this effect in both backgrounds also shows that PolB is active to some extent in both originless and wild-type strains.

Repair of stalled replication forks may be performed through several pathways (as described in section 1.4.3), which may involve simply remodelling and resetting of the fork, or may involve recombination between homologous chromosomes. Depending on the exact scenario, double-strand breaks may be generated, which may also be repaired through recombination. In originless strains, which rely on recombination for replication of their genome, a small increase in stalled forks could result in increased efficiency of genome replication; thus explaining the small growth increase seen in low-dose aphidicolin treatments.

Based on the results described in Chapter 6, a suggested model is that PolB is the dominant DNA polymerase for origin-initiated replication, while PolD is the dominant DNA polymerase used in

recombination-initiated replication. In strains with wild-type origins, aphidicolin treatment causes many stalled forks which need to be repaired; a process in which Hel308 is an important player. High numbers of stalled forks places a burden on the cells which reduces growth rate. Meanwhile, in originless strains, PolB is less active, but still active at some level. A small number of stalled forks causes a beneficial increase in recombination rate; at higher levels this still causes inconvenience to growth rate, but at a lower level than the wild-type due to the lower activity of PolB in these cells.

8. References

Albers, S., Ashmore, J., Pollard, T., Spang, A., & Zhou, J. (2022). Origin of eukaryotes: What can be learned from the first successfully isolated Asgard archaeon. *Fac Rev*, *11*, 3. https://doi.org/10.12703/r-01-000005

Allers, T., Barak, S., Liddell, S., Wardell, K., & Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Appl Environ Microbiol*, *76*(6), 1759-1769. https://doi.org/10.1128/AEM.02670-09

Allers, T., Ngo, H. P., Mevarech, M., & Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon Haloferax volcanii based on the leuB and trpA genes. *Appl Environ Microbiol*, *70*(2), 943-953. https://doi.org/10.1128/aem.70.2.943-953.2004

Anand, R., Buechelmaier, E., Belan, O., Newton, M., Vancevska, A., Kaczmarczyk, A., Takaki, T., Rueda, D. S., Powell, S. N., & Boulton, S. J. (2022). HELQ is a dual-function DSB repair enzyme modulated by RPA and RAD51. *Nature*, *601*(7892), 268-273. https://doi.org/10.1038/s41586-021-04261-0

Anand, R. P., Lovett, S. T., & Haber, J. E. (2013). Break-induced DNA replication. *Cold Spring Harb Perspect Biol*, *5*(12), a010397. https://doi.org/10.1101/cshperspect.a010397

Antón J, Meseguer I, Rodríguez-Valera F. Production of an Extracellular Polysaccharide by *Haloferax mediterranei*. *Appl Environ Microbiol*. 1988 Oct;54(10):2381-6. doi: 10.1128/aem.54.10.2381-2386.1988. PMID: 16347749; PMCID: PMC204266.

Aoki, M., Ehara, M., Saito, Y., Yoshioka, H., Miyazaki, M., Saito, Y., Miyashita, A., Kawakami, S., Yamaguchi, T., Ohashi, A., Nunoura, T., Takai, K., & Imachi, H. (2014). A long-term cultivation of an anaerobic methane-oxidizing microbial community from deep-sea methane-seep sediment using a continuous-flow bioreactor. *PLoS ONE*, *9*(8), e105356. https://doi.org/10.1371/journal.pone.0105356

Arakawa, K., & Tomita, M. (2007). The GC skew index: a measure of genomic compositional asymmetry and the degree of replicational selection. *Evol Bioinform Online*, *3*, 159-168.

Arakawa, T., Yamaguchi, R., Tokunaga, H., & Tokunaga, M. (2017). Unique Features of Halophilic Proteins. *Current Protein & Peptide Science*, *18*(1), 65-71. https://doi.org/10.2174/1389203717666160617111140

Ausiannikava, D., & Allers, T. (2017). Diversity of DNA Replication in the Archaea. *Genes (Basel)*, *8*(2). https://doi.org/10.3390/genes8020056

Ausiannikava, D., Mitchell, L., Marriott, H., Smith, V., Hawkins, M., Makarova, K. S., Koonin, E. V., Nieduszynski, C. A., & Allers, T. (2018). Evolution of Genome Architecture in Archaea: Spontaneous Generation of a New Chromosome in *Haloferax volcanii*. *Molecular Biology and Evolution*, *35*(8), 1855–1868. https://doi.org/https://doi.org/10.1093/molbev/msy075

Babski, J., Haas, K. A., Näther-Schindler, D., Pfeiffer, F., Förstner, K. U., Hammelmann, M., Hilker, R., Becker, A., Sharma, C. M., Marchfelder, A., & Soppa, J. (2016). Genome-wide identification of transcriptional start sites in the haloarchaeon Haloferax volcanii based on differential RNA-Seq (dRNA-Seq). *BMC Genomics*, *17*(1), 629. https://doi.org/10.1186/s12864-016-2920-y

Babu, M. M. (2016). The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem Soc Trans*, *44*(5), 1185-1200. https://doi.org/10.1042/bst20160172

Ball, L. J., Kühne, R., Schneider-Mergener, J., & Oschkinat, H. (2005). Recognition of Proline-Rich Motifs by Protein–Protein-Interaction Domains. *Angewandte Chemie International Edition*, *44*(19), 2852-2869. https://doi.org/https://doi.org/10.1002/anie.200400618

Baranovskiy, A. G., Babayeva, N. D., Suwa, Y., Gu, J., Pavlov, Y. I., & Tahirov, T. H. (2014). Structural basis for inhibition of DNA replication by aphidicolin. *Nucleic Acids Res*, *42*(22), 14013-14021. https://doi.org/10.1093/nar/gku1209

Bauer, R. J., Begley, M. T., & Trakselis, M. A. (2012). Kinetics and fidelity of polymerization by DNA polymerase III from Sulfolobus solfataricus. *Biochemistry*, *51*(9), 1996-2007. https://doi.org/10.1021/bi201799a

Baum, D. A., & Baum, B. (2014). An inside-out origin for the eukaryotic cell. *BMC Biology*, *12*(1), 76. https://doi.org/10.1186/s12915-014-0076-2

Berquist, B. R., DasSarma, P., & DasSarma, S. (2007). Essential and non-essential DNA replication genes in the model halophilic Archaeon, Halobacterium sp. NRC-1. *BMC Genet*, *8*, 31. https://doi.org/10.1186/1471-2156-8-31

Besse, A., Peduzzi, J., Rebuffat, S., & Carrε-Mlouka, A. (2015). Antimicrobial peptides and proteins in the face of extremes: Lessons from archaeocins. *Biochimie*, *118*, 344-355. https://doi.org/https://doi.org/10.1016/j.biochi.2015.06.004

Bidle, K. A., Kirkland, P. A., Nannen, J. L., & Maupin-Furlow, J. A. (2008). Proteomic analysis of Haloferax volcanii reveals salinity-mediated regulation of the stress response protein PspA. *Microbiology (Reading)*, *154*(Pt 5), 1436-1443. https://doi.org/10.1099/mic.0.2007/015586-0

Bitan-Banin, G., Ortenberg, R., & Mevarech, M. (2003). Development of a gene knockout system for the halophilic archaeon *Haloferax volcanii* by use of the *pyrE* gene. *J Bacteriol*, *185*(3), 772-778. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=12533452

Blackwood, J. K., Rzechorzek, N. J., Bray, S. M., Maman, J. D., Pellegrini, L., & Robinson, N. P. (2013). End-resection at DNA double-strand breaks in the three domains of life. *Biochem Soc Trans*, *41*(1), 314-320. https://doi.org/10.1042/bst20120307

Bohall, P. B., & Bell, S. D. (2021). Phenotypic Characterization of Sulfolobus islandicus Strains Lacking the B-Family DNA Polymerases PolB2 and PolB3 Individually and in Combination [Original Research]. *Frontiers in Microbiology*, *12*. https://doi.org/10.3389/fmicb.2021.666974

Braun, F., Thomalla, L., van der Does, C., Quax, T. E. F., Allers, T., Kaever, V., & Albers, S. V. (2019). Cyclic nucleotides in archaea: Cyclic di-AMP in the archaeon Haloferax volcanii and its putative role. *Microbiologyopen*, *8*(9), e00829. https://doi.org/10.1002/mbo3.829

Brent, M. M., Anand, R., & Marmorstein, R. (2008). Structural Basis for DNA Recognition by FoxO1 and Its Regulation by Posttranslational Modification. *Structure*, *16*(9), 1407-1416. https://doi.org/https://doi.org/10.1016/j.str.2008.06.013

Bruert, S., Allers, T., Spohn, G., & Soppa, J. (2006). Regulated Polyploidy in Halophilic Archaea. *PLoS ONE*, *1*(1), e92. https://doi.org/doi:10.1371/journal.pone.0000092

Brunk, C. F., & Martin, W. F. (2019). Archaeal Histone Contributions to the Origin of Eukaryotes. *Trends in Microbiology*, *27*(8), 703-714. https://doi.org/https://doi.org/10.1016/j.tim.2019.04.002

Büttner, K., Nehring, S., & Hopfner, K. P. (2007). Structural basis for DNA duplex separation by a superfamily-2 helicase. *Nat Struct Mol Biol*, *14*(7), 647-652. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17558417

Byrd, A. K., & Raney, K. D. (2012). Superfamily 2 helicases. *Front Biosci (Landmark Ed)*, *17*(6), 2070-2088. https://doi.org/10.2741/4038

Byun, T. S., Pacek, M., Yee, M. C., Walter, J. C., & Cimprich, K. A. (2005). Functional uncoupling of MCM helicase and DNA polymerase activities activates the ATR-dependent checkpoint. *Genes Dev*, *19*(9), 1040-1052. https://doi.org/10.1101/gad.1301205

Calmann, M. A., & Marinus, M. G. (2004). MutS inhibits RecA-mediated strand exchange with platinated DNA substrates. *Proceedings of the National Academy of Sciences*, *101*(39), 14174-14179. https://doi.org/doi:10.1073/pnas.0406104101

Cannan, W. J., & Pederson, D. S. (2016). Mechanisms and Consequences of Double-Strand DNA Break Formation in Chromatin. *J Cell Physiol*, *231*(1), 3-14. https://doi.org/10.1002/jcp.25048

Charlesworth, J., Kimyon, O., Manefield, M., Beloe, C. J., & Burns, B. P. (2020). Archaea join the conversation: detection of AHL-like activity across a range of archaeal isolates. *FEMS Microbiol Lett*, *367*(16). https://doi.org/10.1093/femsle/fnaa123

Chatterjee, N., & Walker, G. C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen*, *58*(5), 235-263. https://doi.org/10.1002/em.22087

Chaudhury, I., Stroik, D. R., & Sobeck, A. (2014). FANCD2-controlled chromatin access of the Fanconi-associated nuclease FAN1 is crucial for the recovery of stalled replication forks. *Mol Cell Biol*, *34*(21), 3939-3954. https://doi.org/10.1128/mcb.00457-14

Chen, Z., Yang, H., & Pavletich, N. P. (2008). Mechanism of homologous recombination from the RecA–ssDNA/dsDNA structures. *Nature*, *453*(7194), 489-494. https://doi.org/10.1038/nature06971

Chimileski, S., Franklin, M. J., & Papke, R. T. (2014). Biofilms formed by the archaeon Haloferax volcanii exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol*, *12*, 65. https://doi.org/10.1186/s12915-014-0065-5

Choi, J. Y., Eoff, R. L., Pence, M. G., Wang, J., Martin, M. V., Kim, E. J., Folkmann, L. M., & Guengerich, F. P. (2011). Roles of the four DNA polymerases of the crenarchaeon Sulfolobus solfataricus and accessory proteins in DNA replication. *J Biol Chem*, *286*(36), 31180-31193. https://doi.org/10.1074/jbc.M111.258038

Chong, J. P., Hayashi, M. K., Simon, M. N., Xu, R. M., & Stillman, B. (2000). A double-hexamer archaeal minichromosome maintenance protein is an ATP-dependent DNA helicase. *Proc Natl Acad Sci U S A*, *97*(4), 1530-1535. https://doi.org/10.1073/pnas.030539597

Cline, S. W., Lam, W. L., Charlebois, R. L., Schalkwyk, L. C., & Doolittle, W. F. (1989). Transformation methods for halophilic archaebacteria. *Can J Microbiol*, *35*(1), 148-152. https://doi.org/10.1139/m89-022

Croteau, D. L., Popuri, V., Opresko, P. L., & Bohr, V. A. (2014). Human RecQ helicases in DNA repair, recombination, and replication. *Annu Rev Biochem*, *83*, 519-552. https://doi.org/10.1146/annurev-biochem-060713-035428

Cubonová, L., Richardson, T., Burkhart, B. W., Kelman, Z., Connolly, B. A., Reeve, J. N., & Santangelo, T. J. (2013). Archaeal DNA polymerase D but not DNA polymerase B is required for genome replication in Thermococcus kodakarensis. *J Bacteriol*, *195*(10), 2322-2328. https://doi.org/10.1128/jb.02037-12

Darwin, K. H., & Hofmann, K. (2010). SAMPyling proteins in archaea. *Trends Biochem Sci*, *35*(6), 348-351. https://doi.org/10.1016/j.tibs.2010.03.003

Dattani, A., Harrison, C., & Allers, T. (2022). Genetic Manipulation of Haloferax Species. In S. Ferreira-Cerca (Ed.), *Archaea: Methods and Protocols* (pp. 33-56). Springer US. https://doi.org/10.1007/978-1-0716-2445-6_3

Dattani, A., Sharon, I., Shtifman-Segal, E., Robinzon, S., Gophna, U., Allers, T., & Altman-Price, N. (2022). Differences in homologous recombination and maintenance of heteropolyploidy between Haloferax volcanii and Haloferax mediterranei. *G3 Genes|Genomes|Genetics*, *13*(4). https://doi.org/10.1093/g3journal/jkac306

de Silva, R. T., Abdul-Halim, M. F., Pittrich, D. A., Brown, H. J., Pohlschroder, M., & Duggin, I. G. (2021). Improved growth and morphological plasticity of Haloferax volcanii. *Microbiology (Reading)*, *167*(2). https://doi.org/10.1099/mic.0.001012

Deem, A., Keszthelyi, A., Blackgrove, T., Vayl, A., Coffey, B., Mathur, R., Chabes, A., & Malkova, A. (2011). Break-induced replication is highly inaccurate. *PLoS Biol*, *9*(2), e1000594. https://doi.org/10.1371/journal.pbio.1000594

Delmas, S., Shunburne, L., Ngo, H. P., & Allers, T. (2009). Mre11-Rad50 promotes rapid repair of DNA damage in the polyploid archaeon Haloferax volcanii by restraining homologous recombination. *PLoS Genet*, *5*(7), e1000552. https://doi.org/10.1371/journal.pgen.1000552

Dombrowski, N., Lee, J. H., Williams, T. A., Offre, P., & Spang, A. (2019). Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiol Lett*, *366*(2). https://doi.org/10.1093/femsle/fnz008

Donnianni, R. A., & Symington, L. S. (2013). Break-induced replication occurs by conservative DNA synthesis. *Proc Natl Acad Sci U S A*, *110*(33), 13475-13480. https://doi.org/10.1073/pnas.1309800110

Eme, L., Spang, A., Lombard, J., Stairs, C. W., & Ettema, T. J. G. (2017). Archaea and the origin of eukaryotes. *Nat Rev Microbiol*. https://doi.org/10.1038/nrmicro.2017.154

Enzmann, F., Mayer, F., Rother, M., & Holtmann, D. (2018). Methanogens: biochemical background and biotechnological applications. *AMB Express*, *8*. https://doi.org/https://doi.org/10.1186/s13568-017-0531-x

Feng, X., Zhang, B., Gao, Z., Xu, R., Liu, X., Ishino, S., Feng, M., Shen, Y., Ishino, Y., & She, Q. (2022). A Well-Conserved Archaeal B-Family Polymerase Functions as an Extender in Translesion Synthesis. *mBio*, *13*(1), e02659-02621. https://doi.org/doi:10.1128/mbio.02659-21

Fernandez-Castillo, R., Rodriguez-Valera, F., Gonzalez-Ramos, J., & Ruiz-Berraquero, F. (1986). Accumulation of Poly (beta-Hydroxybutyrate) by Halobacteria. *Appl Environ Microbiol*, *51*(1), 214-216. https://doi.org/10.1128/aem.51.1.214-216.1986

Frank, A. C., & Lobry, J. R. (2000). Oriloc: prediction of replication boundaries in unannotated bacterial chromosomes. *Bioinformatics*, *16*(6), 560-561. https://doi.org/10.1093/bioinformatics/16.6.560

Fujikane, R., Komori, K., Shinagawa, H., & Ishino, Y. (2005). Identification of a novel helicase activity unwinding branched DNAs from the hyperthermophilic archaeon, Pyrococcus furiosus. *J Biol Chem*, *280*(13), 12351-12358. https://doi.org/10.1074/jbc.M413417200

Fujikane, R., Shinagawa, H., & Ishino, Y. (2006). The archaeal Hjm helicase has recQ-like functions, and may be involved in repair of stalled replication fork. *Genes to Cells*, *11*(2), 99-110. https://doi.org/https://doi.org/10.1111/j.1365-2443.2006.00925.x

Gehring, A. M., Astling, D. P., Matsumi, R., Burkhart, B. W., Kelman, Z., Reeve, J. N., Jones, K. L., & Santangelo, T. J. (2017). Genome Replication in Thermococcus kodakarensis Independent of Cdc6 and an Origin of Replication. *Front Microbiol*, *8*, 2084. https://doi.org/10.3389/fmicb.2017.02084

Geier, G. E., & Modrich, P. (1979). Recognition sequence of the dam methylase of Escherichia coli K12 and mode of cleavage of Dpn I endonuclease. *J Biol Chem*, *254*(4), 1408-1413.

Giani, M., & Martínez-Espinosa, R. M. (2020). Carotenoids as a Protection Mechanism against Oxidative Stress in Haloferax mediterranei. *Antioxidants*, *9*(11), 1060. https://www.mdpi.com/2076-3921/9/11/1060

Giroux, X., & MacNeill, S. A. (2015). Molecular Genetic Methods to Study DNA Replication Protein Function in Haloferax volcanii, A Model Archaeal Organism. In S. Vengrova & J. Dalgaard (Eds.), *DNA Replication: Methods and Protocols* (pp. 187-218). Springer New York. https://doi.org/10.1007/978-1-4939-2596-4_13

Gong, P., Lei, P., Wang, S., Zeng, A., & Lou, H. (2020). Post-Translational Modifications Aid Archaeal Survival. *Biomolecules*, *10*(4). https://doi.org/10.3390/biom10040584

Grasso, S., & Tell, G. (2014). Base excision repair in Archaea: back to the future in DNA repair. *DNA Repair (Amst)*, *21*, 148-157. https://doi.org/10.1016/j.dnarep.2014.05.006

Greci, M. D., Dooher, J. D., & Bell, S. D. (2022). The combined DNA and RNA synthetic capabilities of archaeal DNA primase facilitate primer hand-off to the replicative DNA polymerase. *Nature Communications*, *13*(1), 433. https://doi.org/10.1038/s41467-022-28093-2

Greenough, L., Kelman, Z., & Gardner, A. F. (2015). The roles of family B and D DNA polymerases in Thermococcus species 9°N Okazaki fragment maturation. *J Biol Chem*, *290*(20), 12514-12522. https://doi.org/10.1074/jbc.M115.638130

Gregor, D., & Pfeifer, F. (2005). In vivo analyses of constitutive and regulated promoters in halophilic archaea. *Microbiology (Reading)*, *151*(Pt 1), 25-33. https://doi.org/10.1099/mic.0.27541-0

Gueguen, Y., Rolland, J. L., Lecompte, O., Azam, P., Le Romancer, G., Flament, D., Raffin, J. P., & Dietrich, J. (2001). Characterization of two DNA polymerases from the hyperthermophilic euryarchaeon Pyrococcus abyssi. *Eur J Biochem*, *268*(22), 5961-5969. https://doi.org/10.1046/j.0014-2956.2001.02550.x

Gupta, R. S., Naushad, S., & Baker, S. (2015). Phylogenomic analyses and molecular signatures for the class Halobacteria and its two major clades: a proposal for division of the class Halobacteria into an emended order Halobacteriales and two new orders, Haloferacales ord. nov. and Natrialbales ord. nov., containing the novel families Haloferacaceae fam. nov. and Natrialbaceae fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, *65*(Pt_3), 1050-1069. https://doi.org/https://doi.org/10.1099/ijs.0.070136-0

Guy, C. P., & Bolt, E. L. (2005). Archaeal Hel308 helicase targets replication forks *in vivo* and *in vitro* and unwinds lagging strands. *Nucleic Acids Res*, *33*(11), 3678-3690. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15994460

Guy, L., & Ettema, T. J. (2011). The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol*, *19*(12), 580-587. https://doi.org/10.1016/j.tim.2011.09.002

Hakem, R. (2008). DNA-damage repair; the good, the bad, and the ugly. *Embo j*, *27*(4), 589-605. https://doi.org/10.1038/emboj.2008.15

Han, J., Lu, Q., Zhou, L., Liu, H., & Xiang, H. (2009). Identification of the polyhydroxyalkanoate (PHA)-specific acetoacetyl coenzyme A reductase among multiple FabG paralogs in Haloarcula hispanica and reconstruction of the PHA biosynthetic pathway in Haloferax volcanii. *Appl Environ Microbiol*, *75*(19), 6168-6175. https://doi.org/10.1128/aem.00938-09

Han, J., Zhang, F., Hou, J., Liu, X., Li, M., Liu, H., Cai, L., Zhang, B., Chen, Y., Zhou, J., Hu, S., & Xiang, H. (2012). Complete genome sequence of the metabolically versatile halophilic archaeon Haloferax mediterranei, a poly(3-hydroxybutyrate-co-3-hydroxyvalerate) producer. *J Bacteriol*, *194*(16), 4463-4464. https://doi.org/10.1128/JB.00880-12

Haque, R. U., Paradisi, F., & Allers, T. (2020). Haloferax volcanii for biotechnology applications: challenges, current state and perspectives. *Appl Microbiol Biotechnol*, *104*(4), 1371-1382. https://doi.org/10.1007/s00253-019-10314-2

Harami, G. M., Gyimesi, M., & Kovács, M. (2013). From keys to bulldozers: expanding roles for winged helix domains in nucleic-acid-binding proteins. *Trends in Biochemical Sciences*, *38*(7), 364-371. https://doi.org/https://doi.org/10.1016/j.tibs.2013.04.006

Harrison, C., & Allers, T. (2022). Progress and Challenges in Archaeal Genetic Manipulation. In S. Ferreira-Cerca (Ed.), *Archaea Methods and Protocols* (1 ed., pp. 25-31). Humana New York. https://doi.org/https://doi.org/10.1007/978-1-0716-2445-6

Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., Robinson, J., Khouri, H., Ren, Q., Lowe, T. M., Maupin-Furlow, J., Pohlschroder, M., Daniels, C., Pfeiffer, F., Allers, T., & Eisen, J. A. (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS ONE*, *5*(3), e9605. https://doi.org/10.1371/journal.pone.0009605

Hatano, T., Palani, S., Papatziamou, D., Salzer, R., Souza, D. P., Tamarit, D., Makwana, M., Potter, A., Haig, A., Xu, W., Townsend, D., Rochester, D., Bellini, D., Hussain, H. M. A., Ettema, T. J. G., Löwe, J., Baum, B., Robinson, N. P., & Balasubramanian, M. (2022). Asgard archaea shed light on the evolutionary origins of the eukaryotic ubiquitin-ESCRT machinery. *Nature Communications*, *13*(1), 3398. https://doi.org/10.1038/s41467-022-30656-2

Hawkins, M., Malla, S., Blythe, M. J., Nieduszynski, C. A., & Allers, T. (2013). Accelerated growth in the absence of DNA replication origins. *Nature*, *503*(7477), 544-547. https://doi.org/10.1038/nature12650

He, M. H., Liu, J. C., Lu, Y. S., Wu, Z. J., Liu, Y. Y., Wu, Z., Peng, J., & Zhou, J. Q. (2019). KEOPS complex promotes homologous recombination via DNA resection. *Nucleic Acids Res*, *47*(11), 5684-5697. https://doi.org/10.1093/nar/gkz228

Henneke, G., Flament, D., Hübscher, U., Querellou, J., & Raffin, J.-P. (2005). The Hyperthermophilic Euryarchaeota Pyrococcus abyssi Likely Requires the Two DNA Polymerases D and B for DNA Replication. *Journal of Molecular Biology*, *350*(1), 53-64. https://doi.org/https://doi.org/10.1016/j.jmb.2005.04.042

Hogrel, G., Lu, Y., Alexandre, N., Bossé, A., Dulermo, R., Ishino, S., Ishino, Y., & Flament, D. (2020). Role of RadA and DNA Polymerases in Recombination-Associated DNA Synthesis in Hyperthermophilic Archaea. *Biomolecules*, *10*(7). https://doi.org/10.3390/biom10071045

Holmes, M. L., & Dyall-Smith, M. L. (1991). Mutations in DNA gyrase result in novobiocin resistance in halophilic archaebacteria. *J Bacteriol*, *173*(2), 642-648. https://doi.org/10.1128/jb.173.2.642-648.1991

Holmes, M. L., & Dyall-Smith, M. L. (2000). Sequence and expression of a halobacterial beta-galactosidase gene. *Mol Microbiol*, *36*(1), 114-122. https://doi.org/10.1046/j.1365-2958.2000.01832.x

Hong, Y., Chu, M., Li, Y., Ni, J., Sheng, D., Hou, G., She, Q., & Shen, Y. (2012). Dissection of the functional domains of an archaeal Holliday junction helicase. *DNA Repair*, *11*(2), 102-111. https://doi.org/https://doi.org/10.1016/j.dnarep.2011.10.009

Howley, P. M., Israel, M. A., Law, M. F., & Martin, M. A. (1979). A rapid method for detecting and mapping homology between heterologous DNAs. Evaluation of polyomavirus genomes. *J Biol Chem*, *254*(11), 4876-4883.

Imachi, H., Nobu, M. K., Nakahara, N., Morono, Y., Ogawara, M., Takaki, Y., Takano, Y., Uematsu, K., Ikuta, T., Ito, M., Matsui, Y., Miyazaki, M., Murata, K., Saito, Y., Sakai, S., Song, C., Tasumi, E., Yamanaka, Y., Yamaguchi, T., . . . Takai, K. (2020). Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature*, *577*(7791), 519-525. https://doi.org/10.1038/s41586-019-1916-6

Ishino, S., & Ishino, Y. (2014). DNA polymerases as useful reagents for biotechnology - the history of developmental research in the field. *Front Microbiol*, *5*, 465. https://doi.org/10.3389/fmicb.2014.00465

Jenkins, T., Northall, S. J., Ptchelkine, D., Lever, R., Cubbon, A., Betts, H., Taresco, V., Cooper, C. D. O., McHugh, P. J., Soultanas, P., & Bolt, E. L. (2021). The HelQ human DNA repair helicase utilizes a PWI-like domain for DNA loading through interaction with RPA, triggering DNA unwinding by the HelQ helicase core. *NAR Cancer*, *3*(1), zcaa043. https://doi.org/10.1093/narcan/zcaa043

Jia, H., Dantuluri, S., Margulies, S., Smith, V., Lever, R., Allers, T., Koh, J., Chen, S., & Maupin-Furlow, J. A. (2023). RecJ3/4-aRNase J form a Ubl-associated nuclease complex functioning in survival against DNA damage in *Haloferax volcanii*. *mBio*, *14*(4), e00852-00823. https://doi.org/doi:10.1128/mbio.00852-23

Johansson, E., & Dixon, N. (2013). Replicative DNA polymerases. *Cold Spring Harb Perspect Biol*, *5*(6). https://doi.org/10.1101/cshperspect.a012799

Jokela, M., Eskelinen, A., Pospiech, H., Rouvinen, J., & Syväoja, J. E. (2004). Characterization of the 3' exonuclease subunit DP1 of Methanococcus jannaschii replicative DNA polymerase D. *Nucleic Acids Res*, *32*(8), 2430-2440. https://doi.org/10.1093/nar/gkh558

Jones, N. T. (2019). *The role of homologous recombination in Haloferax volcanii* . PhD thesis, University of Nottingham.

Kavitha, P., Lipton, A. P., Sarika, A. R., & Aishwarya, M. S. (2011). Growth characteristics and halocin production by a new isolate, Haloferax volcanii KPS1 from Kovalam solar saltern (India).

Research Journal of Biological Sciences, *6*, 257-262. https://doi.org/10.3923/rjbsci.2011.257.262

Kazlauskas, D., Krupovic, M., Guglielmini, J., Forterre, P., & Venclovas, Č. (2020). Diversity and evolution of B-family DNA polymerases. *Nucleic Acids Res*, *48*(18), 10142-10156. https://doi.org/10.1093/nar/gkaa760

Kelman, L. M., & Kelman, Z. (2014). Archaeal DNA Replication. In B. L. Bassler (Ed.), *Annual Review of Genetics, Vol 48* (Vol. 48, pp. 71-97). https://doi.org/10.1146/annurev-genet-120213-092148

Killelea, T., Ghosh, S., Tan, S. S., Heslop, P., Firbank, S. J., Kool, E. T., & Connolly, B. A. (2010). Probing the interaction of archaeal DNA polymerases with deaminated bases using X-ray crystallography and non-hydrogen bonding isosteric base analogues. *Biochemistry*, *49*(27), 5772-5781. https://doi.org/10.1021/bi100421r

Koller, M. (2019). Polyhydroxyalkanoate Biosynthesis at the Edge of Water Activitiy-Haloarchaea as Biopolyester Factories. *Bioengineering (Basel)*, *6*(2). https://doi.org/10.3390/bioengineering6020034

Komori, K., Hidaka, M., Horiuchi, T., Fujikane, R., Shinagawa, H., & Ishino, Y. (2004). Cooperation of the N-terminal Helicase and C-terminal Endonuclease Activities of Archaeal Hef Protein in Processing Stalled Replication Forks*. *Journal of Biological Chemistry*, *279*(51), 53175-53185. https://doi.org/https://doi.org/10.1074/jbc.M409243200

Komori, K., Sakae, S., Fujikane, R., Morikawa, K., Shinagawa, H., & Ishino, Y. (2000). Biochemical characterization of the hjc holliday junction resolvase of Pyrococcus furiosus. *Nucleic Acids Res*, *28*(22), 4544-4551. https://doi.org/10.1093/nar/28.22.4544

Kondratick, C. M., Washington, M. T., & Spies, M. (2021). Making choices: DNA replication fork recovery mechanisms. *Seminars in Cell & Developmental Biology*, *113*, 27-37. https://doi.org/https://doi.org/10.1016/j.semcdb.2020.10.001

Kramara, J., Osia, B., & Malkova, A. (2018). Break-Induced Replication: The Where, The Why, and The How. *Trends Genet*, *34*(7), 518-531. https://doi.org/10.1016/j.tig.2018.04.002

Kristensen, T., Cherian, R. M., Gray, F., & MacNeill, S. (2014). The haloarchaeal MCM proteins: bioinformatic analysis and targeted mutagenesis of the β7-β8 and β9-β10 hairpin loops and conserved zinc binding domain cysteines [Original Research]. *Frontiers in Microbiology*, *5*. https://doi.org/10.3389/fmicb.2014.00123

Kropp, H. M., Betz, K., Wirth, J., Diederichs, K., & Marx, A. (2017). Crystal structures of ternary complexes of archaeal B-family DNA polymerases. *PLoS ONE*, *12*(12), e0188005. https://doi.org/10.1371/journal.pone.0188005

Kucukyildirim, S., Behringer, M., Williams, E. M., Doak, T. G., & Lynch, M. (2020). Estimation of the Genome-Wide Mutation Rate and Spectrum in the Archaeal Species Haloferax volcanii. *Genetics*, *215*(4), 1107-1116. https://doi.org/10.1534/genetics.120.303299

Kushida, T., Narumi, I., Ishino, S., Ishino, Y., Fujiwara, S., Imanaka, T., & Higashibata, H. (2019). Pol B, a Family B DNA Polymerase, in Thermococcus kodakarensis is Important for DNA Repair, but not DNA Replication. *Microbes Environ*, *34*(3), 316-326. https://doi.org/10.1264/jsme2.ME19075

Laass, S., Monzon, V. A., Kliemt, J., Hammelmann, M., Pfeiffer, F., Förstner, K. U., & Soppa, J. (2019). Characterization of the transcriptome of Haloferax volcanii, grown under four different conditions, with mixed RNA-Seq. *PLoS ONE*, *14*(4), e0215986. https://doi.org/10.1371/journal.pone.0215986

Lam, W. L., & Doolittle, W. F. (1989). Shuttle vectors for the archaebacterium Halobacterium volcanii. *Proceedings of the National Academy of Sciences*, *86*(14), 5478-5482. https://doi.org/doi:10.1073/pnas.86.14.5478

Lapkouski, M., Panjikar, S., Janscak, P., Smatanova, I. K., Carey, J., Ettrich, R., & Csefalvay, E. (2009). Structure of the motor subunit of type I restriction-modification complex EcoR124I. *Nat Struct Mol Biol*, *16*(1), 94-95. https://doi.org/10.1038/nsmb.1523

Large, A., Stamme, C., Lange, C., Duan, Z., Allers, T., Soppa, J., & Lund, P. A. (2007). Characterization of a tightly controlled promoter of the halophilic archaeon *Haloferax volcanii* and its use in the analysis of the essential *cct1* gene. *Mol Microbiol*, *66*(5), 1092-1106. https://doi.org/10.1111/j.1365-2958.2007.05980.x

Le Breton, M., Henneke, G., Norais, C., Flament, D., Myllykallio, H., Querellou, J., & Raffin, J. P. (2007). The heterodimeric primase from the euryarchaeon Pyrococcus abyssi: a multifunctional enzyme for initiation and repair? *J Mol Biol*, *374*(5), 1172-1185. https://doi.org/10.1016/j.jmb.2007.10.015

Leigh, J. A., Albers, S.-V., Atomi, H., & Allers, T. (2011). Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiology Reviews*, *35*(4), 577-608. https://doi.org/10.1111/j.1574-6976.2011.00265.x

Lestini, R., Duan, Z., & Allers, T. (2010). The archaeal Xpf/Mus81/FANCM homolog Hef and the Holliday junction resolvase Hjc define alternative pathways that are essential for cell viability in Haloferax volcanii. *DNA Repair*, *9*(9), 994-1002. https://doi.org/https://doi.org/10.1016/j.dnarep.2010.06.012

Lestini, R., Laptenok, S. P., Kühn, J., Hink, M. A., Schanne-Klein, M.-C., Liebl, U., & Myllykallio, H. (2013). Intracellular dynamics of archaeal FANCM homologue Hef in response to halted DNA replication. *Nucleic Acids Research*, *41*(22), 10358-10370. https://doi.org/10.1093/nar/gkt816

Lever, R. (2020). *Genetic and Biochemical Analysis of the Hel308 Helicase in the Archaeon Haloferax volcanii.* PhD thesis, University of Nottingham.

Lever, Rebecca J., Simmons, E., Gamble-Milner, R., Buckley, Ryan J., Harrison, C., Parkes, Ashley J., Mitchell, L., Gausden, Jacob A., Škulj, S., Bertoša, B., Bolt, Edward L., & Allers, T. (2023). Archaeal Hel308 suppresses recombination through a catalytic switch that controls DNA annealing. *Nucleic Acids Research*, *51*(16), 8563-8574. https://doi.org/10.1093/nar/gkad572

Levin, M. K., Gurjar, M., & Patel, S. S. (2005). A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat Struct Mol Biol*, *12*(5), 429-435. https://doi.org/10.1038/nsmb920

Li, S., Wang, H., Jehi, S., Li, J., Liu, S., Wang, Z., Truong, L., Chiba, T., Wang, Z., & Wu, X. (2021). PIF1 helicase promotes break-induced replication in mammalian cells. *Embo j*, *40*(8), e104509. https://doi.org/10.15252/embj.2020104509

Li, Z., Lu, S., Hou, G., Ma, X., Sheng, D., Ni, J., & Shen, Y. (2008). Hjm/Hel308A DNA Helicase from *Sulfolobus tokodaii* Promotes Replication Fork Regression and Interacts with Hjc Endonuclease In Vitro. *Journal of Bacteriology*, *190*(8), 3006-3017. https://doi.org/doi:10.1128/jb.01662-07

Liao, H., Ji, F., Helleday, T., & Ying, S. (2018). Mechanisms for stalled replication fork stabilization: new targets for synthetic lethality strategies in cancer treatments. *EMBO Rep*, *19*(9). https://doi.org/10.15252/embr.201846263

Liao, Y., Ithurbide, S., Evenhuis, C., Löwe, J., & Duggin, I. G. (2021). Cell division in the archaeon Haloferax volcanii relies on two FtsZ proteins with distinct functions in division ring assembly and constriction. *Nat Microbiol*, *6*(5), 594-605. https://doi.org/10.1038/s41564-021-00894-z

Lieber, M. R. (2010). The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu Rev Biochem*, *79*, 181-211. https://doi.org/10.1146/annurev.biochem.052308.093131

Lin, Z., Kong, H., Nei, M., & Ma, H. (2006). Origins and evolution of the recA/RAD51 gene family: evidence for ancient gene duplication and endosymbiotic gene transfer. *Proc Natl Acad Sci U S A*, *103*(27), 10328-10333. https://doi.org/10.1073/pnas.0604232103

Lindahl, T. (1993). Instability and decay of the primary structure of DNA. *Nature*, *362*(6422), 709-715. https://doi.org/10.1038/362709a0

Ludt, K., & Soppa, J. (2018). Influence of Origin Recognition Complex Proteins on the Copy Numbers of Three Chromosomes in Haloferax volcanii. *J Bacteriol*, *10*(200). https://doi.org/doi:10.1128/JB.00161-18

MacNeill, S. A. (2009). The haloarchaeal chromosome replication machinery. *Biochem Soc Trans*, *37*(Pt 1), 108-113. https://doi.org/10.1042/bst0370108

Maier, L. K., Stachler, A. E., Brendel, J., Stoll, B., Fischer, S., Haas, K. A., Schwarz, T. S., Alkhnbashi, O. S., Sharma, K., Urlaub, H., Backofen, R., Gophna, U., & Marchfelder, A. (2019). The nuts and bolts of the Haloferax CRISPR-Cas system I-B. *RNA Biol*, *16*(4), 469-480. https://doi.org/10.1080/15476286.2018.1460994

Makarova, K. S., Krupovic, M., & Koonin, E. V. (2014). Evolution of replicative DNA polymerases in archaea and their contributions to the eukaryotic replication machinery. *Front Microbiol*, *5*, 354. https://doi.org/10.3389/fmicb.2014.00354

Malkova, A., & Ira, G. (2013). Break-induced replication: functions and molecular mechanism. *Curr Opin Genet Dev*, *23*(3), 271-279. https://doi.org/10.1016/j.gde.2013.05.007

Malkova, A., Naylor, M. L., Yamaguchi, M., Ira, G., & Haber, J. E. (2005). RAD51-dependent break-induced replication differs in kinetics and checkpoint responses from RAD51-mediated gene conversion. *Mol Cell Biol*, *25*(3), 933-944. https://doi.org/10.1128/mcb.25.3.933-944.2005

Marinsek, N., Barry, E. R., Makarova, K. S., Dionne, I., Koonin, E. V., & Bell, S. D. (2006). GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep*, *7*(5), 539-545. https://doi.org/10.1038/sj.embor.7400649

Marinus, M. G. (2000). Recombination is essential for viability of an Escherichia coli dam (DNA adenine methyltransferase) mutant. *J Bacteriol*, *182*(2), 463-468. https://doi.org/10.1128/jb.182.2.463-468.2000

Marinus, M. G. (2010). DNA methylation and mutator genes in Escherichia coli K-12. *Mutat Res*, *705*(2), 71-76. https://doi.org/10.1016/j.mrrev.2010.05.001

Marriott, H. (2018). *Genome architecture and DNA replication in Haloferax volcanii* (Publication Number 50190) University of Nottingham].

Martin, I. V., & MacNeill, S. A. (2002). ATP-dependent DNA ligases. *Genome Biol*, *3*(4), Reviews3005. https://doi.org/10.1186/gb-2002-3-4-reviews3005

Matsumi, R., Manabe, K., Fukui, T., Atomi, H., & Imanaka, T. (2007). Disruption of a Sugar Transporter Gene Cluster in a Hyperthermophilic Archaeon Using a Host-Marker System Based on Antibiotic Resistance. *J Bacteriol*, *189*(7), 2683-2691. https://doi.org/10.1128/JB.01692-06

Mayanagi, K., Ishino, S., Shirai, T., Oyama, T., Kiyonari, S., Kohda, D., Morikawa, K., & Ishino, Y. (2018). Direct visualization of DNA baton pass between replication factors bound to PCNA. *Sci Rep*, *8*(1), 16209. https://doi.org/10.1038/s41598-018-34176-2

McCaffrey, R., St Johnston, D., & González-Reyes, A. (2006). Drosophila mus301/spindle-C encodes a helicase with an essential role in double-strand DNA break repair and meiotic progression. *Genetics*, *174*(3), 1273-1285. https://doi.org/10.1534/genetics.106.058289

McCulloch, B. (2021). *DNA replication in growth conditions that mimic the natural habitat of Haloferax volcanii* (Publication Number 65463) University of Nottingham.

McGlynn, P., Lloyd, R. G., & Marians, K. J. (2001). Formation of Holliday junctions by regression of nascent DNA in intermediates containing stalled replication forks: RecG stimulates regression even when the DNA is negatively supercoiled. *Proc Natl Acad Sci U S A*, *98*(15), 8235-8240. https://doi.org/10.1073/pnas.121007798

McMillan, L. J., Hwang, S., Farah, R. E., Koh, J., Chen, S., & Maupin-Furlow, J. A. (2018). Multiplex quantitative SILAC for analysis of archaeal proteomes: a case study of oxidative stress responses. *Environ Microbiol*, *20*(1), 385-401. https://doi.org/10.1111/1462-2920.14014

Merino, N., Aronson, H. S., Bojanova, D. P., Feyhl-Buska, J., Wong, M. L., Zhang, S., & Giovannelli, D. (2019). Living at the Extremes: Extremophiles and the Limits of Life in a Planetary Context [Review]. *Frontiers in Microbiology*, *10*(780). https://doi.org/10.3389/fmicb.2019.00780

Meselson, M., & Stahl, F. W. (1958). The Replication of DNA in *Escherichia coli*. *Proc Natl Acad Sci U S A*, *44*(7), 671-682. https://doi.org/10.1073/pnas.44.7.671

Mitra, R., Xu, T., Xiang, H., & Han, J. (2020). Current developments on polyhydroxyalkanoates synthesis by using halophiles as a promising cell factory. *Microb Cell Fact*, *19*(1), 86. https://doi.org/10.1186/s12934-020-01342-z

Mullakhanbhai, M. F., & Larsen, H. (1975). Halobacterium volcanii spec. nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch Microbiol*, *104*(3), 207-214. https://doi.org/10.1007/bf00447326

Nagata, M., Ishino, S., Yamagami, T., Ogino, H., Simons, J. R., Kanai, T., Atomi, H., & Ishino, Y. (2017). The Cdc45/RecJ-like protein forms a complex with GINS and MCM, and is important for DNA replication in *Thermococcus kodakarensis*. *Nucleic Acids Res*, *45*(18), 10693-10705. https://doi.org/10.1093/nar/gkx740

Nagata, M., Ishino, S., Yamagami, T., Simons, J. R., Kanai, T., Atomi, H., & Ishino, Y. (2017). Possible function of the second RecJ-like protein in stalled replication fork repair by interacting with Hef. *Sci Rep*, *7*(1), 16949. https://doi.org/10.1038/s41598-017-17306-0

Naor, A., Lapierre, P., Mevarech, M., Papke, R. T., & Gophna, U. (2012). Low species barriers in halophilic archaea and the formation of recombinant hybrids. *Curr Biol*, *22*(15), 1444-1448. https://doi.org/10.1016/j.cub.2012.05.056

Naor, A., Lazary, R., Barzel, A., Papke, R. T., & Gophna, U. (2011). In vivo characterization of the homing endonuclease within the polB gene in the halophilic archaeon *Haloferax volcanii*. *PLoS ONE*, *6*(1), e15833. https://doi.org/10.1371/journal.pone.0015833

Naor, A., Thiaville, P. C., Altman-Price, N., Cohen-Or, I., Allers, T., de Crécy-Lagard, V., & Gophna, U. (2012). A genetic investigation of the KEOPS complex in halophilic Archaea. *PLoS ONE*, *7*(8), e43013. https://doi.org/10.1371/journal.pone.0043013

Nguyen-Hieu, T., Khelaifia, S., Aboudharam, G., & Drancourt, M. (2012). Methanogenic archaea in subgingival sites: a review. *Acta Pathologica, Microbiologica, et Immunologica Scandinavica*, *121*(6), 467-477. https://doi.org/https://doi.org/10.1111/apm.12015

Norais, C., Hawkins, M., Hartman, A. L., Eisen, J. A., Myllykallio, M., & Allers, T. (2007). Genetic and physical mapping of DNA replication origins in *Haloferax volcanii*. *PLoS Genet*, *18*(5), e77. https://doi.org/10.1371/journal.pgen.0030077

Ortenberg, R., Rozenblatt-Rosen, O., & Mevarech, M. (2000). The extremely halophilic archaeon *Haloferax volcanii* has two very different dihydrofolate reductases. *Mol Microbiol*, *35*(6), 1493-1505. https://doi.org/10.1046/j.1365-2958.2000.01815.x

Oyama, T., Ishino, S., Fujino, S., Ogino, H., Shirai, T., Mayanagi, K., Saito, M., Nagasawa, N., Ishino, Y., & Morikawa, K. (2011). Architectures of archaeal GINS complexes, essential DNA replication initiation factors. *BMC Biol*, *9*, 28. https://doi.org/10.1186/1741-7007-9-28

Oyama, T., Ishino, S., Shirai, T., Yamagami, T., Nagata, M., Ogino, H., Kusunoki, M., & Ishino, Y. (2016). Atomic structure of an archaeal GAN suggests its dual roles as an exonuclease in DNA repair and a CMG component in DNA replication. *Nucleic Acids Res*, *44*(19), 9505-9517. https://doi.org/10.1093/nar/gkw789

Oyama, T., Oka, H., Mayanagi, K., Shirai, T., Matoba, K., Fujikane, R., Ishino, Y., & Morikawa, K. (2009). Atomic structures and functional implications of the archaeal RecQ-like helicase Hjm. *BMC Struct Biol*, *9*, 2. https://doi.org/10.1186/1472-6807-9-2

Pacek, M., Tutter, A. V., Kubota, Y., Takisawa, H., & Walter, J. C. (2006). Localization of MCM2-7, Cdc45, and GINS to the Site of DNA Unwinding during Eukaryotic DNA Replication. *Molecular Cell*, *21*(4), 581-587. https://doi.org/https://doi.org/10.1016/j.molcel.2006.01.030

Paul, S., Bag, S. K., Das, S., Harvill, E. T., & Dutta, C. (2008). Molecular signature of hypersaline adaptation: insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol*, *9*(4), R70. https://doi.org/10.1186/gb-2008-9-4-r70

Pérez-Arnaiz, P., Dattani, A., Smith, V., & Allers, T. (2020). *Haloferax volcanii*—a model archaeon for studying DNA replication and repair. *Open Biology*, *10*(12). https://doi.org/10.1098/rsob.200293

Pfeiffer, F., & Dyall-Smith, M. (2021). Open Issues for Protein Function Assignment in *Haloferax volcanii* and Other Halophilic Archaea. *Genes (Basel)*, *12*(7). https://doi.org/10.3390/genes12070963

Pohlschroder, M., & Schulze, S. (2019). Microbe of the Month: *Haloferax volcanii*. *Trends in Microbiology*, *27*(1), 86-87. https://doi.org/10.1016/j.tim.2018.10.004

Qi, Q., Ito, Y., Yoshimatsu, K., & Fujiwara, T. (2016). Transcriptional regulation of dimethyl sulfoxide respiration in a haloarchaeon, *Haloferax volcanii*. *Extremophiles*, *20*(1), 27-36. https://doi.org/10.1007/s00792-015-0794-6

Rados, T., Andre, K., Cerletti, M., & Bisson, A. (2023). A sweet new set of inducible and constitutive promoters in *Haloferax volcanii*. *Front Microbiol*, *14*, 1204876. https://doi.org/10.3389/fmicb.2023.1204876

Rafiq, M., Hassan, N., Rehman, M., Hayat, M., Nadeem, G., Hassan, F., Iqbal, N., Ali, H., Zada, S., Kang, Y., Sajjad, W., & Jamal, M. (2023). Challenges and Approaches of Culturing the Unculturable Archaea. *Biology (Basel)*, *12*(12). https://doi.org/10.3390/biology12121499

Reuter, C. J., & Maupin-Furlow, J. (2004). Analysis of proteasome-dependent proteolysis in *Haloferax volcanii* cells, using short-lived green fluorescent proteins. *Appl Environ Microbiol*, *70*(12), 7530-7538. http://www.ncbi.nlm.nih.gov/pubmed/15574956

Richards, J. D., Johnson, K. A., Liu, H., McRobbie, A. M., McMahon, S., Oke, M., Carter, L., Naismith, J. H., & White, M. F. (2008). Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J Biol Chem*, *283*(8), 5118-5126. https://doi.org/10.1074/jbc.M707548200

Robu, M. E., Inman, R. B., & Cox, M. M. (2001). RecA protein promotes the regression of stalled replication forks in vitro. *Proc Natl Acad Sci U S A*, *98*(15), 8211-8218. https://doi.org/10.1073/pnas.131022698

Rodrigues-Oliveira, T., Wollweber, F., Ponce-Toledo, R. I., Xu, J., Rittmann, S. K. M. R., Klingl, A., Pilhofer, M., & Schleper, C. (2023). Actin cytoskeleton and complex cell architecture in an Asgard archaeon. *Nature*, *613*(7943), 332-339. https://doi.org/10.1038/s41586-022-05550-y

Rogozin, I. B., Makarova, K. S., Pavlov, Y. I., & Koonin, E. V. (2008). A highly conserved family of inactivated archaeal B family DNA polymerases. *Biol Direct*, *3*, 32. https://doi.org/10.1186/1745-6150-3-32

Rosenshine, I., Tchelet, R., & Mevarech, M. (1989). The mechanism of DNA transfer in the mating system of an archaebacterium. *Science*, *245*(4924), 1387-1389. https://doi.org/10.1126/science.2818746

Rouillon, C., Henneke, G., Flament, D., Querellou, J., & Raffin, J.-P. (2007). DNA Polymerase Switching on Homotrimeric PCNA at the Replication Fork of the Euryarchaea *Pyrococcus abyssi*. *Journal of Molecular Biology*, *369*(2), 343-355. https://doi.org/https://doi.org/10.1016/j.jmb.2007.03.054

Saini, N., Ramakrishnan, S., Elango, R., Ayyar, S., Zhang, Y., Deem, A., Ira, G., Haber, J. E., Lobachev, K. S., & Malkova, A. (2013). Migrating bubble during break-induced replication drives conservative DNA synthesis. *Nature*, *502*(7471), 389-392. https://doi.org/10.1038/nature12584

Samson, R. Y., Abeyrathne, P. D., & Bell, S. D. (2015). Mechanism of Archaeal MCM Helicase Recruitment to DNA Replication Origins. *Mol Cell*, *61*(2), 287-296. https://doi.org/10.1016/j.molcel.2015.12.005

Samson, R. Y., Xu, Y., Gadelha, C., Stone, T. A., Faqiri, J. N., Li, D., Qin, N., Pu, F., Liang, Y. X., She, Q., & Bell, S. D. (2013). Specificity and function of archaeal DNA replication initiator proteins. *Cell Rep*, *3*(2), 485-496. https://doi.org/10.1016/j.celrep.2013.01.002

Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*, *74*(12), 5463-5467. https://doi.org/10.1073/pnas.74.12.5463

Santangelo, T. J., Cubonová, L., Matsumi, R., Atomi, H., Imanaka, T., & Reeve, J. N. (2008). Polarity in archaeal operon transcription in *Thermococcus kodakaraensis*. *J Bacteriol*, *190*(6), 2244-2248. https://doi.org/10.1128/jb.01811-07

Sarmiento, F., Mrázek, J., & Whitman, W. B. (2013). Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*. *Proc Natl Acad Sci U S A*, *110*(12), 4726-4731. https://doi.org/10.1073/pnas.1220225110

Sauguet, L., Raia, P., Henneke, G., & Delarue, M. (2016). Shared active site architecture between archaeal PolD and multi-subunit RNA polymerases revealed by X-ray crystallography. *Nat Commun*, *7*, 12227. https://doi.org/10.1038/ncomms12227

Shimmin, L. C., & Dennis, P. P. (1996). Conserved sequence elements involved in regulation of ribosomal protein gene expression in halophilic archaea. *J Bacteriol*, *178*(15), 4737-4741. https://doi.org/10.1128/jb.178.15.4737-4741.1996

Shivanand, P., & Mugeraya, G. (2011). Halophilic bacteria and their compatible solutes – osmoregulation and potential applications. *Current Science*, *100*(10), 1516-1521.

Skowyra, A., & MacNeill, S. A. (2012). Identification of essential and non-essential single-stranded DNA-binding proteins in a model archaeal organism. *Nucleic Acids Res*, *40*(3), 1077-1090. https://doi.org/10.1093/nar/gkr838

Smith, V. (2021). *Atypical DNA replication in Haloferax volcanii in the absence of replication origins* (Publication Number 69014) University of Nottingham.

Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., van Eijk, R., Schleper, C., Guy, L., & Ettema, T. J. G. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature*, *521*(7551), 173-179. https://doi.org/10.1038/nature14447

Srivastav, R., Sharma, R., Tandon, S., & Tandon, C. (2019). Role of DHH superfamily proteins in nucleic acids metabolism and stress tolerance in prokaryotes and eukaryotes. *Macromolecules*, *127*, 66-75. https://doi.org/https://doi.org/10.1016/j.ijbiomac.2018.12.123

Stachler, A. E., & Marchfelder, A. (2016). Gene Repression in Haloarchaea Using the CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas I-B System. *J Biol Chem*, *291*(29), 15226-15242. https://doi.org/10.1074/jbc.M116.724062

Stroud, A., Liddell, S., & Allers, T. (2012). Genetic and Biochemical Identification of a Novel Single-Stranded DNA-Binding Complex in Haloferax volcanii. *Front Microbiol*, *3*, 224. https://doi.org/10.3389/fmicb.2012.00224

Sun, M., Feng, X., Liu, Z., Han, W., Liang, Y. X., & She, Q. (2018). An Orc1/Cdc6 ortholog functions as a key regulator in the DNA damage response in Archaea. *Nucleic Acids Res*, *46*(13), 6697-6711. https://doi.org/10.1093/nar/gky487

Tafel, A. A., Wu, L., & McHugh, P. J. (2011). Human HEL308 localizes to damaged replication forks and unwinds lagging strand structures. *J Biol Chem*, *286*(18), 15832-15840. https://doi.org/10.1074/jbc.M111.228189

Takashima, N., Ishino, S., Oki, K., Takafuji, M., Yamagami, T., Matsuo, R., Mayanagi, K., & Ishino, Y. (2019). Elucidating functions of DP1 and DP2 subunits from the *Thermococcus kodakarensis* family D DNA polymerase. *Extremophiles*, *23*(1), 161-172. https://doi.org/10.1007/s00792-018-1070-3

Tang, N., Wen, W., Liu, Z., Xiong, X., & Wu, Y. (2023). HELQ as a DNA helicase: Its novel role in normal cell function and tumorigenesis (Review). *Oncol Rep*, *50*(6). https://doi.org/10.3892/or.2023.8657

Thomas, C. M., Desmond-Le Quéméner, E., Gribaldo, S., & Borrel, G. (2022). Factors shaping the abundance and diversity of the gut archaeome across the animal kingdom. *Nature Communications*, *13*(1), 3358. https://doi.org/10.1038/s41467-022-31038-4

Timpson, L. M., Liliensiek, A. K., Alsafadi, D., Cassidy, J., Sharkey, M. A., Liddell, S., Allers, T., & Paradisi, F. (2013). A comparison of two novel alcohol dehydrogenase enzymes (ADH1 and ADH2) from the extreme halophile *Haloferax volcanii*. *Appl Microbiol Biotechnol*, *97*(1), 195-203. https://doi.org/10.1007/s00253-012-4074-4

Tori, K., Kimizu, M., Ishino, S., & Ishino, Y. (2007). DNA polymerases BI and D from the hyperthermophilic archaeon *Pyrococcus furiosus* both bind to proliferating cell nuclear antigen with their C-terminal PIP-box motifs. *J Bacteriol*, *189*(15), 5652-5657. https://doi.org/10.1128/jb.00073-07

Turkowyd, B., Schreiber, S., Wörtz, J., Segal, E. S., Mevarech, M., Duggin, I. G., Marchfelder, A., & Endesfelder, U. (2020). Establishing Live-Cell Single-Molecule Localization Microscopy Imaging and Single-Particle Tracking in the Archaeon *Haloferax volcanii*. *Front Microbiol*, *11*, 583010. https://doi.org/10.3389/fmicb.2020.583010

Vashee, S., Cvetic, C., Lu, W., Simancek, P., Kelly, T. J., & Walter, J. C. (2003). Sequence-independent DNA binding and replication initiation by the human origin recognition complex. *Genes Dev*, *17*(15), 1894-1908. https://doi.org/10.1101/gad.1084203

Vesela, E., Chroma, K., Turi, Z., & Mistrik, M. (2017). Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. *Biomolecules*, *7*(1). https://doi.org/10.3390/biom7010019

von Kügelgen, A., Alva, V., & Bharat, T. A. M. (2021). Complete atomic structure of a native archaeal cell surface. *Cell Reports*, *37*(8), 110052. https://doi.org/https://doi.org/10.1016/j.celrep.2021.110052

Wardell, K., Haldenby, S., Jones, N., Liddell, S., Ngo, G. H. P., & Allers, T. (2017). RadB acts in homologous recombination in the archaeon *Haloferax volcanii*, consistent with a role as recombination mediator. *DNA Repair (Amst)*, *55*, 7-16. https://doi.org/10.1016/j.dnarep.2017.04.005

Wardle, J., Burgers, P. M., Cann, I. K., Darley, K., Heslop, P., Johansson, E., Lin, L. J., McGlynn, P., Sanvoisin, J., Stith, C. M., & Connolly, B. A. (2008). Uracil recognition by replicative DNA polymerases is limited to the archaea, not occurring with bacteria and eukarya. *Nucleic Acids Res*, *36*(3), 705-711. https://doi.org/10.1093/nar/gkm1023

Wellington, M., & Rustchenko, E. (2005). 5-Fluoro-orotic acid induces chromosome alterations in Candida albicans. *Yeast*, *22*(1), 57-70. https://doi.org/10.1002/yea.1191

Wendoloski, D., Ferrer, C., & Dyall-Smith, M. L. (2001). A new simvastatin (mevinolin)-resistance marker from Haloarcula hispanica and a new *Haloferax volcanii* strain cured of plasmid pHV2. *Microbiology (Reading)*, *147*(Pt 4), 959-964. https://doi.org/10.1099/00221287-147-4-959

White, M. F., & Allers, T. (2018). DNA repair in the archaea—an emerging picture. *FEMS Microbiology Reviews*, *42*(4), 514-526. https://doi.org/10.1093/femsre/fuy020

Winker, S., & Woese, C. R. (1991). A definition of the domains Archaea, Bacteria and Eucarya in terms of small subunit ribosomal RNA characteristics. *Syst Appl Microbiol*, *14*(4), 305-310. https://doi.org/10.1016/s0723-2020(11)80303-6

Winter, J. A., Christofi, P., Morroll, S., & Bunting, K. A. (2009). The crystal structure of *Haloferax volcanii* proliferating cell nuclear antigen reveals unique surface charge characteristics due to halophilic adaptation. *BMC Struct Biol*, *9*, 55. https://doi.org/1472-6807-9-55 [pii]

Winter, K., Born, J., & Pfeifer, F. (2018). Interaction of Haloarchaeal Gas Vesicle Proteins Determined by Split-GFP. *Front Microbiol*, *9*, 1897. https://doi.org/10.3389/fmicb.2018.01897

Woese, C. R., & Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A*, *74*(11), 5088-5090. http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=270744

Woese, C. R., Kandler, O., & Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A*, *87*(12), 4576-4579. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC54159/pdf/pnas01037-0173.pdf

Woodman, I. L., & Bolt, E. L. (2009). Molecular biology of Hel308 helicase in archaea. *Biochem Soc Trans*, *37*(Pt 1), 74-78. https://doi.org/BST0370074 [pii] 10.1042/BST0370074

Woodman, I. L., & Bolt, E. L. (2011). Winged helix domains with unknown function in Hel308 and related helicases. *Biochem Soc Trans*, *39*(1), 140-144. https://doi.org/10.1042/BST0390140

Woodman, I. L., Brammer, K., & Bolt, E. L. (2011). Physical interaction between archaeal DNA repair helicase Hel308 and Replication Protein A (RPA). *DNA Repair (Amst)*, *10*(3), 306-313. https://doi.org/10.1016/j.dnarep.2010.12.001

Woods, W. G., & Dyall-Smith, M. L. (1997). Construction and analysis of a recombination-deficient (radA) mutant of *Haloferax volcanii*. *Mol Microbiol*, *23*(4), 791-797. https://doi.org/10.1046/j.1365-2958.1997.2651626.x

Wu, Z., Liu, H., Liu, J., Liu, X., & Xiang, H. (2012). Diversity and evolution of multiple orc/cdc6-adjacent replication origins in haloarchaea. *BMC Genomics*, *13*, 478. https://doi.org/10.1186/1471-2164-13-478

Wu, Z., Liu, J., Yang, H., Liu, H., & Xiang, H. (2014). Multiple replication origins with diverse control mechanisms in *Haloarcula hispanica*. *Nucleic Acids Res*, *42*(4), 2282-2294. https://doi.org/10.1093/nar/gkt1214

Wurch, L., Giannone, R. J., Belisle, B. S., Swift, C., Utturkar, S., Hettich, R. L., Reysenbach, A.-L., & Podar, M. (2016). Genomics-informed isolation and characterization of a symbiotic Nanoarchaeota system from a terrestrial geothermal environment. *Nature Communications*, *7*(1), 12115. https://doi.org/10.1038/ncomms12115

Xu, Y., Gristwood, T., Hodgson, B., Trinidad, J. C., Albers, S.-V., & Bell, S. D. (2016). Archaeal orthologs of Cdc45 and GINS form a stable complex that stimulates the helicase activity of MCM. *Proceedings of the National Academy of Sciences*, *113*(47), 13390-13395. https://doi.org/doi:10.1073/pnas.1613825113

Yang, H. B., Wu, Z. F., Liu, J. F., Liu, X. Q., Wang, L., Cai, S. F., & Xiang, H. (2015). Activation of a dormant replication origin is essential for *Haloferax mediterranei* lacking the primary origins [Article]. *Nature Communications*, *6*, 11, Article 8321. https://doi.org/10.1038/ncomms9321

Yeeles, J. T., Poli, J., Marians, K. J., & Pasero, P. (2013). Rescuing stalled or damaged replication forks. *Cold Spring Harb Perspect Biol*, *5*(5), a012815. https://doi.org/10.1101/cshperspect.a012815

Zerulla, K., Chimileski, S., Näther, D., Gophna, U., Papke, T., & Soppa, J. (2014). DNA as a Phosphate Storage Polymer and the Alternative Advantages of Polyploidy for Growth or Survival. *PLoS ONE*, *9*(4), e94819. https://doi.org/ doi:10.1371/journal.pone.0094819

Zhang, C., Tian, B., Li, S., Ao, X., Dalgaard, K., Gökce, S., Liang, Y., & She, Q. (2013). Genetic manipulation in *Sulfolobus islandicus* and functional analysis of DNA repair genes. *Biochem Soc Trans*, *41*(1), 405-410. https://doi.org/10.1042/bst20120285

Zhang, R., & Zhang, C. T. (2005). Identification of replication origins in archaeal genomes based on the Z-curve method. *Archaea*, *1*(5), 335-346. https://doi.org/10.1155/2005/509646

Zhao, A., Gray, F. C., & MacNeill, S. A. (2006). ATP- and NAD+-dependent DNA ligases share an essential function in the halophilic archaeon *Haloferax volcanii*. *Mol Microbiol*, *59*(3), 743-752. https://doi.org/10.1111/j.1365-2958.2005.04975.x

Zheng, T., Huang, Q., Zhang, C., Ni, J., She, Q., & Shen, Y. (2012). Development of a Simvastatin Selection Marker for a Hyperthermophilic Acidophile, *Sulfolobus islandicus*. *Appl Environ Microbiol*, *78*(2). https://doi.org/https://doi.org/10.1128/AEM.06095-11

Zhong, D. (2015). Electron transfer mechanisms of DNA repair by photolyase. *Annu Rev Phys Chem*, *66*, 691-715. https://doi.org/10.1146/annurev-physchem-040513-103631